

aer

**THE ALBERTA JOURNAL OF
EDUCATIONAL RESEARCH**

VOLUME XXX

NUMBER 2

JUNE 1984

**PUBLISHED BY
THE UNIVERSITY OF ALBERTA • EDMONTON**

THE ALBERTA JOURNAL OF EDUCATIONAL RESEARCH

Published in March, June, September, December by the
Faculty of Education, The University of Alberta

A quarterly journal devoted to the dissemination, criticism, interpretation and encouragement of all forms of systematic enquiry into education and fields related to or associated with education.

EDITOR: C. S. Bumbarger

CONSULTING EDITORS

D. C. Bjorkquist
University of Minnesota

P. E. Bryant
University of Oxford

B. M. Chung
Chungbuk University,
Republic of Korea

A. Côté
University of Ottawa

D. Fitzgerald
The University of New England

D. B. Harrison
The University of Calgary

N. L. Hersom
The University of Saskatchewan

M. Laferrière
McGill University

T. Morrison
University of Manitoba

J. Richardson
University of New Brunswick

P. Robinson
Atlantic Institute of Education

T. Rogers
University of British Columbia

R. E. Traub
Ontario Institute for
Studies in Education

K. Watson
University of Reading

R. J. Wilson
Queen's University

The views expressed and the accuracy of the statements made are the responsibility of the individual authors. The editor is solely responsible for the editorial comments.

AJER gratefully acknowledges support from the Social Sciences and Humanities Research Council of Canada and from the Alberta Advisory Committee for Educational Studies.

AJER is indexed in the *Canadian Education Index*, *Current Contents/Social and Behavioral Sciences*, *Social Science Citation Index*, *ERIC/Current Index to Journals in Education*, and *Alberta Education Index*; appropriate articles are abstracted in *Educational Administration Abstracts*, *Psychological Abstracts*, *Sociology of Education Abstracts*, *Language Behavior Abstracts*, *Cognitive Development Abstracts*, *Research into Higher Education Abstracts*, *School Organization and Management Abstracts*, and *Multicultural Education Abstracts*.

The subscription rate is \$12.00 per year; single copies are \$3.50 each. Please make cheques payable to *The Alberta Journal of Educational Research*. All back issues are available; rates supplied on request. Claims for undelivered copies must be received within three months of publication.

Address all communications and manuscript submissions to *The Alberta Journal of Educational Research*, Faculty of Education, Publication Services, 4-116 Education North, The University of Alberta, Edmonton, Canada, T6G 2G5.

SECOND CLASS MAIL REGISTRATION NUMBER 1436



The Alberta Journal of Educational Research

Volume XXX, Number 2

June, 1984

CONTENTS

The Effect of Homogeneous vs. Heterogeneous Matching-Item Format on Test Performance and Reliability	91
<i>D. E. Allison</i>	

A Detailed Analysis of Group Differences on the California Short-Form Test of Mental Maturity Between 1956 and 1977	95
<i>U. Dash and T. Maguire</i>	

An Empirical Analysis of Loans by School Libraries	104
<i>R. Houle and C. Montmarquette</i>	

Serving on the School Board: A Political Apprenticeship for Higher Office	115
<i>N. Robinson and C. Stacey</i>	

The Relationship Between Simultaneous-Successive Processing and Academic Achievement	126
<i>F. M. Merritt and R. S. McCallum</i>	

ESSAY REVIEWS

Logo and the Development of General Problem-Solving Skills	133
<i>L. R. Krasnor and J. O. Mitterer</i>	

Self-fulfilling Prophecies in the Classroom: Dead End or Promising Beginning?	145
<i>V. S. Smead</i>	

FACULTY OF EDUCATION
The University of Alberta

DONALD E. ALLISON

University of British Columbia

The Effect of Homogeneous vs. Heterogeneous Matching-Item Format on Test Performance and Reliability

Two forms of a general science matching-item test were prepared. Both forms consisted of three matching exercises. Each of the exercises in Form A contained homogeneous items devoted to a single topic while the exercises in Form B contained a heterogeneous arrangement of items drawn at random from the topics in Form A. When the test was administered to 316 sixth grade students, analysis of the data revealed that, while there was no significant difference between the reliability of the two forms, scores on the heterogeneous form of the test were significantly higher than those on the homogeneous form and that this difference was independent of any interaction with the examinee's sex or intelligence.

Although they do not cite empirical evidence to support their contention, authors of measurement books continue to recommend that matching exercises should be homogeneous (Brown, 1983; Hills, 1981; Hopkins & Antes, 1978; Lien, 1980; Mehrens & Lehmann, 1975; Thorndike & Hagen, 1977). Homogeneity, in this case, means stimuli and responses that are restricted to single concepts, classifications, or areas, so that a homogeneous matching-item test would not likely include persons, places, and things in a single exercise. For the most part, these authors maintain that homogeneous matching exercises require students to demonstrate their knowledge of the subject content by making careful discriminations in selecting the correct responses. On the other hand, exercises which are not homogeneous provide test-wise students with clues which allow them to eliminate responses which have no logical connection to the stimulus, thus enabling them to select the correct responses from a much shorter and easier list.

This study was designed to make an empirical comparison between two formats, homogeneous and heterogeneous, of a matching-item test. Since any difference be-

Donald Allison (PhD, Southern California) is an Associate Professor and Coordinator of the Measurement, Evaluation and Research Methodology area of the Department of Educational Psychology and Special Education, The University of British Columbia. His current research and publications are focused on classroom testing and test construction.

tween the difficulty and reliability of the homogeneous and heterogeneous forms of the test might be due to an advantage that a heterogeneous matching exercise might give to more able and/or test-wise students, less experienced and sophisticated elementary school students were chosen as subjects for this investigation.

It was hypothesized that there is no significant difference in test performance due to (1) matching-item format or (2) any possible interaction between matching-item format, intelligence and sex, or any combination of these three variables, and (3) that there is no significant difference in test reliability due to matching-item format.

Method

The Instrument

Two forms of a sixth grade general science matching-item test were constructed. Form A consisted of three homogeneous, 10-item, 11-option, matching exercises. The first exercise was concerned with the classification of animals, the second with the solar system, and the third with earth science. Each of the exercises was placed on a separate page with the options arranged in alphabetical order. Form B also consisted of three, 10-item, 11-option, matching exercises, but the items and their matching options were drawn at random from Form A. Form B exercises, then, were heterogeneous, containing items from all three topics. As with Form A, each exercise was placed on a separate page with the options in alphabetical order. The maximum possible score for both Forms A and B was 30.

Procedure

The subjects were 316 sixth grade students enrolled in 15 classes in a suburban school district. Forms A and B were distributed alternately within each classroom, so that when all the students had been tested, 156 had completed Form A and 160 Form B.

From intelligence test scores obtained from school records, the mean IQ of these students was found to be 112.37 and the median 113.63. When students were classified as high- or low-IQ by dividing the intelligence test scores at the median, there were 159 high-IQ students and 157 low-IQ students.

Results

The means and standard deviations of the matching-item test scores, when students were classified by intelligence (IQ), sex, and test format, are presented in Table 1. This table shows that the mean score for the total group of 316 students was 18.87 with a standard deviation of 6.06.

The significance of these data was assessed by means of a 2 (IQ) X 2 (sex) X 2 (format) analysis of variance. This analysis is summarized in Table 2.

The results show that the main effect due to format was significant, $p < .01$, causing rejection of the first hypothesis that there would be no difference in test performance due to matching-item format. Table 1 reveals that scores were higher on the heterogeneous format (Form B) than on the homogeneous format (Form A).

No hypotheses related to the other two main effects had been stated. The significant effects ($p < .01$) due to intelligence and sex indicated that the high-IQ students did better on the matching-item test than the low-IQ students and that the science content of the test was easier for the boys than for the girls.

TABLE 1
MEANS AND STANDARD DEVIATIONS OF MATCHING-ITEM
TEST SCORES: STUDENTS CLASSIFIED BY IQ, SEX, AND FORMAT

Variable	Level	<u>n</u>	Mean	<u>SD</u>
IQ	High	159	21.81	5.04
	Low	157	15.89	5.55
Sex	Boys	157	19.87	5.88
	Girls	159	17.88	6.10
Form	A	156	16.67	5.45
	B	160	21.01	5.87
Combined		316	18.87	6.06

TABLE 2
SUMMARY OF THE ANALYSIS OF VARIANCE OF MATCHING-ITEM
TEST SCORES BY INTELLIGENCE, SEX, AND FORMAT

Source of Variation	<u>df</u>	Mean Square	<u>F</u>
IQ	1	2656.973	116.361*
Sex	1	270.862	11.862*
Format	1	1505.178	65.918*
IQ X Sex	1	24.100	1.055
IQ X Format	1	7.059	0.309
Sex X Format	1	0.086	0.004
IQ X Sex X Format	1	10.966	0.480
Within Cells	308	22.834	
Total	315		

*p < .01

As predicted in the second hypothesis, Table 2 also shows that none of the interactions between IQ, sex, and format were significant, $p < .05$.

Kuder-Richardson formula 20 reliabilities were calculated as .777 and .845 for Forms A and B of the test. The difference between these two coefficients, when assessed by the procedure suggested by Feldt (1969) was not found to be significant, $W(155,159) = 1.439$, $p > .05$, thus supporting the third hypothesis, that there is no difference in test reliability due to matching-item format.

Discussion

These results show that even when the subject matter content is identical, the format of a matching-item test influences test performance, and this effect is independent of any interaction with either the sex or intelligence of the examinees. While there was no significant difference in the reliability, scores on the heterogeneous form of this test were significantly higher than those on the homogeneous form, leading to a conclusion that the difference was due to the matching-item format rather than knowledge of the content.

Since scores on an achievement test should be a function of content mastery rather than item format, it can also be concluded that measurement specialists are correct in emphasizing the importance of homogeneity in the design and construction of matching-item exercises.

References

- Brown, F. G. (1983). *Principles of educational and psychological testing* (3rd ed.). New York: Holt, Rinehart and Winston.
- Feldt, L. S. (1969). A test of the hypothesis that Cronbach's alpha or Kuder-Richardson coefficient twenty is the same for two tests. *Psychometrika*, 34, 363-373.
- Hills, J. R. (1981). *Measurement and evaluation in the classroom* (2nd ed.). Columbus, OH: Charles E. Merrill.
- Hopkins, C. D., & Antes, R. L. (1978). *Classroom measurement and evaluation*. Itasca, IL: F. E. Peacock.
- Lien, A. J. (1980). *Measurement and evaluation of learning* (4th ed.). Dubuque, IA: Wm. C. Brown.
- Mehrens, W. A., & Lehmann, I. J. (1975). *Measurement and evaluation in education and psychology* (2nd ed.). New York: Holt, Rinehart and Winston.
- Thorndike, R. L., & Hagen, E. P. (1977). *Measurement and evaluation in psychology and education* (4th ed.). New York: John Wiley and Sons.

UDAYA DASH

Utkal University

and

THOMAS MAGUIRE

University of Alberta

A Detailed Analysis of Group Differences on the California Short-Form Test of Mental Maturity Between 1956 and 1977

The California Short Form Test of Mental Maturity was administered to 3,443 third grade students in 1956, and to 4,378 third grade students in 1977. Differences in factorial structure were examined, as were differences in ability level between the two groups on those factors which appear to be measuring consistent abilities. Of the 98 items, 64 appeared to behave with factorial consistency in the two administrations. Consistent items were related to seven components including two kinds of Sensing Right and Left, Spatial Relationships, Similarities, Numerical Reasoning, Vocabulary and Verbal Mediation. Differences in ability level favored the 1977 group on Sensing Right and Left II, Numerical Reasoning and Verbal Mediation. The 1956 group had a higher mean than the 1977 group on Vocabulary. Results were interpreted in terms of Jensen's ability levels and in terms of specific changes in curriculum and children's television.

There have been a number of reports showing that educational test scores have declined over the past few decades (Harnischfeger & Wiley, 1976).¹ Closer examination reveals that the phenomenon is more complex than was first thought. During the 1940s and through to the mid-'60s, scores seemed to increase fairly steadily, with declines occurring in the past two decades. The declines have tended to be greater in the upper grades than in the elementary grades and there seemed to be differential effects related to student gender.

Udaya Dash is a lecturer in the Department of Psychology, Utkal University, India. He is interested in measurement, applied statistics, and cognitive development.

Thomas Maguire is a professor in the Department of Educational Psychology, University of Alberta. His interests are educational measurement, applied statistics, and research methodology.

Reports on score decline at the elementary school level are not unequivocal. As Harnischfeger and Wiley note, there has been little or no decline in Iowa Tests of Basic Skill scores at the third grade, and there has actually been a modest increase for the Comprehensive Tests of Basic Skills at grades 2, 3, and 4. On the other hand, National Assessment of Educational Progress scores have shown decreases in science for 9- and 13-year-old students during the early seventies.

In Canada, studies over 40 years (Hedges, 1977) and over 21 years (Clarke, Nyberg, & Worth, 1977) have shown the same mixture of results, with declines being noted among older children and no differences or modest increases occurring with younger children.

Although less well-studied than achievement, school-related ability measures provide a similar pattern. In high schools, declining scores have been noted in the Scholastic Aptitude Test, the American College Testing Program, the Preliminary Scholastic Aptitude Test, and the Minnesota Scholastic Aptitude Test. At grade 3, Nyberg and Blackmore (1981) report an increase of about one-half of a standard deviation in raw scores on the California Short Form Test of Mental Maturity.

There have been many speculations concerning the cause for changes in aptitude and achievement scores, amongst which the cumulative effect of TV, retention of proportionately more students in high school, and changes in curriculum have been most prominent. Unfortunately, interpretations have been made difficult by the omnibus nature of the tests used. It seems likely that investigation into the nature of aptitude and achievement differences would be more profitable if attention were paid to differential changes in specific abilities, or specific achievements at the subtest level, or even at the item level. For example, one might expect TV to influence negatively some areas such as reading comprehension, but science knowledge or knowledge of current affairs might be changed positively. Changes in curriculum ought to manifest themselves in a fairly specific fashion, whereas greater school retention rates may lead to more general changes. Thus to begin to understand the nature of the phenomenon, it is necessary to look for differences in specific areas.

Objectives

The purpose of the present study was to examine specific changes in educationally related abilities for two groups of third grade students measured in 1956 and 1977 using the California Short Form Test of Mental Maturity (CSTM)—Primary S. The samples were made up of all third grade students in the Edmonton Public Schools who were present on the testing days in those two years. More specifically the objectives were:

1. To examine differences in factorial structure of item responses between 1977 and students in 1956 on the CSTM.
2. To examine differences in ability level between the two groups on those factors which appear to be measuring consistent abilities.

Method

Sample

The sample consisted of 7,821 Edmonton Public School Board Grade 3 students: 3,443 children were tested in 1956 and 4,378 children were tested in 1977. All children were tested on a large battery of cognitive tests for the project described by Clarke, Nyberg and Worth (1977). The present paper is an extended analysis of some of the data used by Clarke et al. (1977), by Nyberg and Blackmore (1981), and by Blackmore (1980).

Test

The California Short Form Test of Mental Maturity (CSTM), now out of print, was a part of the larger parent test called the California Test of Mental Maturity. The CSTM consists of 98 items grouped into 7 subtests (Sensing Left and Right, Manipulation of Areas, Similarities, Inference, Number Series, Numerical Quantity, and Verbal Concepts) which are used to obtain a general intellectual maturity measure in a one-period group test. The test, its reliability, validity, administration, and scoring procedures are reported in the test manual (Sullivan, Clark, & Tiegs, 1953).

According to Clarke et al. (1977), teachers, supervisors in elementary education from the school system, and experts from the University of Alberta were asked to identify items which would be inappropriate for the 1977 class. In the case of the CSTM no such item was reported. The CSTM makes provision for calculating an IQ score based on the total score for the test. Although time limits are imposed on the students, the test is basically designed as a power test. In Table 1, a detailed description of each of the subtests is given. The subtests are grouped according to what Sullivan et al. call, “major factors involved in intelligence or mental capacity.” These are: spatial relations, logical reasoning, numerical reasoning, and verbal concepts.

TABLE 1

DESCRIPTION OF THE CALIFORNIA SHORT-FORM
TEST OF MENTAL MATURITY

Spatial Relationships—Orientation in space and the use of spatial relationships

Test 1: Sensing Right and Left

The first part of the test consists of 6 pictures of girls and boys and the students are instructed to mark particular appendages such as, “Put a mark on the girl’s left arm.” The second part of the test consists of four double pictures of hands or feet in various positions and the students are told to “Put a mark on each right hand or foot.”

Test 2: Manipulation of Areas

The test consists of 12 picture items intended to measure the students’ ability to use spatial imagery to manipulate different forms. Each item consists of a row of four drawings. The student is told that the “first drawing is among the other drawings but is turned around or turned over. Find it and put a mark on it.”

Logical Reasoning

Test 3: Similarities

This test consists of a total of 12 picture situations. The first two pictures are alike in some ways (e.g., monkey, horse). The pupil determines the nature of this likeness and finds another picture among the other three (tea pot, broom, mouse,) which is similar to the first two.

Test 4: Inference

The test is made up of 12 picture situations. The examiner reads a few short sentences and then asks the students to make an inference. For example, in one situation there is a picture of two boys holding fish. The examiner says, “Look at the two boys. Bill caught more fish than Ned. Put a mark on Bill. He is the boy with the most fish.”

Numerical Reasoning

Test 5: Number Series

There are 12 picture situations which require the pupil to understand lightest, heaviest, most, etc., and number series. Six of the situations consist of pictures (e.g., bicycle, plane, car)

and the student is required to choose the one that falls at the extreme on some characteristic (e.g., put a mark on the thing that can go the fastest). In the other six situations there are five boxes with different numbers of circles marked on them to form a number series. The student is told, "One box is wrong in each row. Put a mark on the box that is wrong."

Test 6: Numerical Quantity

There are 12 picture problems. The student is given a numerical situation and is asked a question. (e.g., "Look at the first box in this row. It has three blocks. If you take one of the blocks away which box will it look like?" Three boxes follow with 2, 5 and 4 blocks in them.)

Test 7: Verbal Concepts

This test consists of 28 items each containing three pictures. The student is given a single word or concept and required to choose the picture that matches the word or concept. For example, the three pictures are: children going upstairs, a merry-go-round, children going downstairs. The examiner says, "Put a mark on those descending."

Data Analysis

The two inter-item correlation matrices, one for the 1956 and the other for the 1977 sample, were analyzed using a principal components analysis. On the basis of the scree test, seven components were retained and rotated to the varimax criterion. The varimax loadings of the 1956 sample were rotated to match the loadings obtained for the 1977 sample using an orthogonal Procrustes solution. The error matrix was used for identifying the items that contributed to differences in factor structure. Any item having a difference of .15 in its loadings on the same component for the two groups was removed. In addition, any item which had differential loadings of .10 or more on *two or more* components was also removed.

Although the sampling distribution of factor loadings is not well explicated, it was felt that with over 300 subjects these criterion values would not only be significant, but important as well. Accordingly, 34 out of the 98 items in the CSTM were removed. The intercorrelation matrix of the remaining 64 items for both groups combined was analyzed and rotated to the varimax criterion. Component scores were calculated for all the subjects. Analysis of variance was performed on the seven component scores between 1956 and 1977 students to compare the achievement differences of both groups, on those parts of the test having a common structure.

Results and Discussion

Differences in Component Structure

Thirty-four items were found to contribute to differences in component structure. In the case of 27, the difference between a rotated component loading for 1956 and a 1977 component loading exceeded .15. (The largest discrepancy was .23.) For 9 of the 27, this occurred on the defining component. Seven of the 34 items were excluded because of two or more discrepancies between .10 and .14. The source of the items which did not fit is shown in Table 2. From that table, it can be seen that there seems to be a disproportionately small number from Right Hands and Feet, Manipulation of Areas, and Logical Inference. In spite of these differences, viewed from an item salience perspective, the fit was remarkably good. On all but 14 items the highest component loading for the 1977 solution agreed with that of the rotated 1956 version.

Factor Structure of the CSTM

Seven components emerged clearly and consistently when separate analyses were performed on the 1956 and 1977 groups, and also when the combined data of both groups on the 64-item portion of the test were analyzed. The first component was defined by items that required subjects to identify the left and right hand or foot of single pictures of boys or girls. This factor was named Sensing Left and Right—I. The second component, which was named Sensing Left and Right—II, was defined by items requiring relatively more complex discriminations between left and right than in the first. They dealt with double picture situations at a relatively more difficult level of operation. The third component was labelled Spatial Relationship, since the items with high loadings required subjects to identify a rotated or an inverted image of the target pattern in the midst of a number of distracting patterns. These items appear to measure subjects' ability to use spatial imagery in manipulating patterns of different forms and in many different posi-

TABLE 2
SOURCES OF ITEMS THAT CONTRIBUTED
TO DIFFERENCES IN FACTOR STRUCTURE

Subtest	Items Contained in Subtest	Items Contributing to Difference	Proportion
Sensing Right & Left			
Body Parts	1 to 6	3,5,6	.5
Right Hand or Foot	7 to 12		0
Manipulation of Areas	1 to 12	6,12	.17
Similarities	1 to 12	2,6,8,10	.33
Inference	1 to 12	7,10	.17
Number Series			
Choose Extreme	1 to 6	3,5,6	.5
Number Series	7 to 12	11,12	.33
Numerical Quality	1 to 12	1,2,3,7,8,9	.5
Verbal Concepts	1 to 28	1,2,3,5,6, 10,12,14,16, 19,21,24	.43
TOTAL	98	34	.35

tions. The fourth component, named Similarities, was defined by items which asked subjects to examine the similarities between two stimulus pictures, and then to identify a picture among the three response pictures that is similar to the stimulus. The fifth component, Numerical Reasoning, consisted of items measuring pupils' ability to make inferences with special reference to quantitative situations and problems. The sixth component, Vocabulary, was defined by items which require the subjects to find pictures of certain objects or animals among a number of other pictures. The items that loaded on the seventh component, Verbal Mediation, test pupils' ability to identify pictures that portray the verbal descriptions specified by the items. (For example, two of the items under this component required subjects to find "something delicious," and find "those descending.") It is interesting that subtest 4 (Inference) did not belong to the space occupied by other items in the test. Indeed when the correlation matrix was examined, none of the 45 correlations between Inference items exceeded .1, and only 5 of the 540 correlations between Inference items and non-Inference items exceeded .1, the largest being .13.

Table 3 presents the clustering of the 64 items on the seven components. The emergence of these seven components in a sample size much larger than the sample (N=700) on which the original structure of the CSTM was based offers support for the validity of the original structure of the CSTM, except of course for the Inference dimension.

TABLE 3
CLUSTERING OF THE 64 ITEMS ON SEVEN COMPONENTS^a

Components	Subtest No.	Item No.
I. Sensing left and right-I	1	1,2,4
II. Sensing left and right-II	1	7,8,9,10
III. Spatial relationship	2	1,2,3,4,5,7,8,9,10,11
IV. Similarities	3	1,3,4,5,7,9,11,12
V. Numerical reasoning	5 6	1,2,4,5,6,7,8 4,5,6,10,11,12
VI. Vocabulary	7	4,7,8,9,11,13,15,18
VII. Verbal mediation	7	17,20,22,23,25,26, 27,28

^a The item Nos. 1,2,3,4,5,6,8,9,11,12 of Subtest No. 4 (Inference) did not load on any component.

Ability Differences Between 1956 and 1977 Groups

Table 4 presents the mean student abilities of 1956 and 1977 Edmonton Grade 3 children based on component scores calculated from the 64 consistent items of the CSTM as well as on the component scores derived from the entire test.

TABLE 4
MEAN STUDENT ABILITIES OF 1956 AND 1977 EDMONTON
GRADE 3 CHILDREN AND F VALUES ON SEVEN COMPONENTS

Factors	1956	1977	F
Sensing left and right-I	.037	-.029	7.04
Sensing left and right-II	-.454	.357	1524.90*
Spatial relationship	-.031	.026	7.11
Similarities	-.023	.020	3.76
Numerical reasoning	-.138	.109	112.13*
Vocabulary	.111	-.084	72.09*
Verbal mediation	-.280	.221	521.32*

Based on 64 items

* Significant <.001

Given the large number of students in the sample, it was decided to use a significance level of .001. This difference between means corresponds to about a tenth of a component score standard deviation. (Significance at the .05 level corresponds to a difference of about one/twenty-fifth of a standard deviation.) By using the higher critical difference, it was hoped that trivial differences could be avoided and as well the overall significance level would be kept below .01.

As may be seen from Table 4, the 1977 group performed at a higher level than the 1956 group on the components labelled Sensing Right and Left II, Numerical Reasoning, and Verbal Mediation. In Vocabulary the 1956 group had a higher mean than the 1977 group.

Several speculations can be made as to the reasons for these differences. Vocabulary as a cognitive skill does not emphasize the processing requirements that

Sensing Left and Right II, Numerical Reasoning, and Verbal Mediation do. In contrast to Vocabulary, the processing requirements of the other three components correspond to Jensen's (1970) Level II ability in the sense that the original input needs to be manipulated, evaluated, and transformed to produce an effective output. On the items defining Vocabulary, the subjects are merely required to point to one of three pictures. The processing requirements are considerably less in Vocabulary for most people.

If curriculum and instruction in 1956 placed more emphasis on knowledge than in 1977, and if a gradual shift toward emphasizing understanding, analysis, and evaluation has taken place since, then the present results would be consistent. The largest difference occurred on Sensing Right and Left II which requires the student to be able to distinguish right and left hands and feet, and to identify them under different orientations. Introduction of manipulative aids to primary school arithmetic may partially account for the observed differences. The kind of skills taught on children's television would also contribute to this difference.

Both Numerical Reasoning and Verbal Mediation involve aural memory and reasoning which are also skills promoted by children's television. The items for Logical Inference (Test 4) which did not contribute to test structure may have exceeded the capability of the children at this level to retain the information in memory and then to carry out the complex analysis required. Even so, the children may have been able to combine intuition with partial information to solve the problems. Whatever the cause, on eleven of the twelve items, the 1977 group had higher means, which supports the contention that aural memory and reasoning skills are slightly higher in the 1977 group.

In their 1981 article, Nyberg and Blackmore note a difference in CSTM mean from 1956 to 1977 of about two-thirds of a standard deviation. Blackmore's (1980) latent trait analysis of the same data showed a difference of about one-half a standard deviation on the ability scale after 25 aberrant items have been removed. In both cases the results favored the 1977 group. The results of the present study suggest that the differences across abilities are not constant and indeed the order is not even consistent. The 1977 students seem to have superior level II processing skills, whereas the 1956 students are favored on vocabulary as measured by picture recognition.

Because the results differentially favor one group in some cases and the other group in other cases, it seems unlikely that global influences of population migration or teacher recruitment are operating. Such influence would seem more likely to produce general differences over many areas favoring a single group. More plausible influences would seem to be specific changes in curriculum and television viewing habits. It is impossible to document all the changes that took place in the 21 years, but closer examination of the data narrows the search. Detailed analysis of the remaining tests in the Clarke, Nyberg, and Worth (1977) battery is being undertaken with the hope of specifying the changes in skills more clearly.

Notes

1. Harnischfeger and Wiley provide an excellent summary of literature in the area. Unless otherwise stated, research cited comes from that review.

References

- Blackmore, D. E. (1980). *A latent trait study of item biachievement differences*. Unpublished doctoral dissertation, University of Alberta, Edmonton, Alberta.
- Clarke, S. C. T., Nyberg, V., & Worth, W. H. (1977). *General report on Edmonton Grade III achievement: 1956-1977 comparisons*. Edmonton, Canada: Alberta Education.
- Harnischfeger, A., & Wiley, D. (1976). Achievement test scores drop. So what? *Educational Researcher*, 5, 5-12.
- Hedges, H. G. (1977). *Achievement in basic skills: A longitudinal evaluation of pupil achievement in language arts and mathematics*. Toronto: Ontario Institute for Studies in Education.
- Jensen, A. R. (1970). Hierarchical theories of mental ability. In W. B. Dockerell (Ed.), *On intelligence*. Toronto: Ontario Institute for Studies in Education.
- Nyberg, V., & Blackmore, D. E. (1981). A longitudinal study of Grade III achievement in Edmonton Public Schools. *Alberta Journal of Educational Research*, 27, 154-159.
- Sullivan, E. T., Clark, W. W., & Tiegs, E. W. *California shortform test of mental maturity, Primary grades 1, 2, 3, 1953 S-Form*. Los Angeles: California Test Bureau.

RACHEL HOULE

and

CLAUDE MONTMARQUETTE

Université de Montréal

An Empirical Analysis of Loans by School Libraries

This empirical study measures the influence of selected factors on loans of printed materials made by a sample of Québec primary and secondary school libraries. A production function approach is chosen with library loans as the output, and library resources and many policy and interaction variables as inputs. The analysis relies on probit and linear regression techniques for the estimation of parameters after a distinction is made between school libraries that make loans and those which do not. The study yields information to help shape policies about the number of loans by school libraries.

Reading is not only one of the basic skills to be acquired but, as well, is considered by both parents and professional educators to be a pedagogical instrument to stimulate the intellectual development of children. School libraries provide both reading materials and a supportive environment and appear, therefore, to constitute a key element in promoting the importance and enjoyment of reading. However, the efficiency of school libraries in developing reading habits is conditional on certain factors among which is the allocation of the meager \$5 and \$20 per student spent annually, from 1972 to 1979, for Québec primary and secondary school libraries respectively.

Rachel Houle has specialized in regional economics and economics of education at the University of Pennsylvania where she obtained her PhD in 1979. Currently Assistant Professor of Economics at Université de Montréal, Miss Houle has produced several scientific articles.

Claude Montmarquette has specialized in public finance and econometrics at the University of Chicago where he obtained his PhD in 1973. Currently Professor of Economics and Director of the Centre de Recherche en Développement Economique (C.R.D.E.), Université de Montréal, Mr. Montmarquette has published several scientific articles.

Library use is a multidimensional concept. School library materials can be consulted on the premises or borrowed; the users can be students, teachers, parents, or other people; there are some inter-library loans; and some libraries also have audio-visual materials. Assuming that book borrowing influences reading habits, the focus of this study is on the loans of printed materials to individuals. Unfortunately, the use of library materials on the premises had to be excluded from the analysis because of the lack of reliable data.

We estimate the role of selected factors on loan policies of school libraries at the elementary and secondary levels by providing answers to the following questions. What part do the quantity and quality of books in the library, the size and composition of personnel, the organizational methods, the physical facilities, and the users' characteristics play in determining whether or not a school library makes loans? For those school libraries making loans, what are the main determinants of the number of loans per pupil made over a school year?

The quantitative analysis performed in the study links the output of school libraries (loans) to their inputs (e.g., number of books or personnel) in a production function framework. The analysis includes numerous explanatory interaction variables and relies on particularly suitable statistical models, namely the probit and multiple regression techniques.

The analysis should yield interesting information to help shape policies designed to increase the number of loans made by school libraries. In a period of severe financial constraint currently affecting education in Québec and elsewhere in Canada, the question of the best use of all available resources is of major importance.

A Model of the Determinants of School Library Loan Services

Contrary to the recent public library literature, the literature on the determinants of school library services at the elementary and secondary levels is very scarce. The few studies on the subject are more descriptive than causal and the discussion is generally qualitatively oriented.¹ The following services associated with school libraries are recognized in this descriptive literature:

1. technical services: selection, acquisition, classification and conservation of materials, and
2. services to users such as information, reference and loans.

To include in the model all services performed by school libraries would require the analysis to be conducted in a multi-output production function framework. Unfortunately, the number of loans is the only output for which reliable quantitative information exists at the present time for a large number of elementary and secondary schools, so that such a framework cannot be employed.

The output variable, the level of which constitutes the explanatory focus, takes two forms in the analysis. The first form of the output variable is a dichotomous variable, that is, a variable taking only two values—whether or not a school library makes loans; the second is the number of loans per pupil made over a school year provided that a school library makes loans.

In a causal model, two sets of explanatory variables are generally distinguished: first, *the policy variables* connected with the human and physical resources and the organization and, second, *the standardizing variables*, which in this case are the users' and the school's characteristics. This second set of variables lies outside the school library operating policies.

TABLE 1

MEAN VALUES AND STANDARD ERRORS
OF SELECTED VARIABLES

Variables	Elementary schools		Secondary schools	
	full sample	with loans	full sample	with loans
Number of schools	877	417	350	283
Number of annual loans (per pupil)		38.0 (30.5) ^a		17.18 (16.1)
Number of Books (per pupil)	16.12 (8.35)	16.5 (9.03)	23.1 (20.3)	22.7 (19.6)
Number of periodicals (per pupil)	.06 (.24)	.07 (.29)	.1 (.6)	.11 (.45)
New acquisitions (% of total books)	.08 (.18)	.10 (.22)	.15 (1.6)	.06 (.08)
Weedings (% of total books)	.02 (.06)	.02 (.06)	.03 (.09)	.03 (.09)
Library personnel per 100 pupils:				
professionals (full-time)	-b	-b	.05	.06
professionals (part-time)	.1	.14	.06	.07
technicians (full-time)	-b	.01	.06	.06
technicians (part-time)	.06	.08	.05	.05
support staff (full-time)	.01	.01	.07	.06
support staff (part-time)	.19	.28	.09	.08
Number of operating hours (per week)	19.7 (10.8)	22.6 (9.2)	30.31 (12.4)	32.2 (10.9)
Library area (in square meters per pupil)	.33 (.48)	.36 (.43)	.62 (.86)	.61 (.85)
Number of carrels (per pupil)	.0032 (.04)	.0045 (.05)	.01 (.02)	.01 (.02)
Number of chairs (per pupil)	.1 (.1)	.12 (.09)	.12 (.11)	.12 (.1)
Number of schools in socioeconomic areas which are				
strongly underprivileged	176	71	49	43
underprivileged	324	146	133	102
privileged	377	200	168	138
Number of schools located in towns				
of less than 10 000 inhabitants	492	221	161	131
of 10 000 to 25 000 inhabitants	122	62	56	42
of 25 000 to 100 000 inhabitants	192	100	87	73
100 000 inhabitants or more	71	34	46	37
Number of teachers in school	15.2 (7.0)	16.1 (7.7)	56.3 (45.2)	62.6 (45.9)
Number of private schools	9	5	56	43

^a Standard Error shown in parentheses

^b Negligible value

TABLE 2
TOTAL ELASTICITIES

Variables	Elementary school		Secondary school	
	probabil- ities of loans ^{a,c}	number of loans	probabil- ities of loans ^{b,c}	number of loans
<u>Size and quality of the stock:</u>				
Number of books (per pupil)	-.042	.454 (.115)	.044	.437 (.157)
Number of periodicals (per pupil)	.023	.148 (.027)	.003	-.076 (.067)
New acquisitions (% of total books)	.194	.045 (.074)	-.81E-07	.026 (.087)
Weedings (% of total books)	.325	.059 (.065)	.176	.041 (.038)
<u>Library personnel:</u>				
Professionals (full-time)	-.022	-.01 (.033)	.034	.016 (.05)
Professionals (part-time)	.013	-.023 (.007)	-.004	.015 (.038)
Technicians (full-time)	.003	.028 (.03)	.012	-.003 (.045)
Technicians (part-time)	.086	.039 (.023)	.165	.078 (.032)
Support staff (full-time)	.359	.079 (.053)	-.017	-.032 (.033)
Support staff (part-time)	.034	-.003 (.017)	.002	.058 (.023)
<u>Library facilities:</u>				
Number of operating hours (per week)	.299	.175 (.11)	.02	.832 (.307)
Library area (in sq. m. per pupil)	.052	-.127 (.071)	-.009	.16 (.096)
Number of booths (per pupil)	-.009	-.027 (.014)	.005	-.059 (.042)
Number of chairs (per pupil)	.184	.132 (.079)	-.019	.146 (.111)
<u>Others:</u>				
Percentage of 1st to 3rd graders	.169	.114 (.091)	-	-
Percentage of students in lower grades	-	-	.062	.187
Percentage of students in professional programs	-	-	.005	-.022 (.048)
Number of teachers	.198	.055 (.086)	.047	-.06 (.107)

^a Computed at mean values of variables and coefficients of regression results reported in tables of the appendix.

^b Computed at mean value of variables minus two standard errors and coefficients of regression results. Owing to a negative value, we substitute the computed value by 1E-06, for 6 variables.

^c Estimated standard errors are not reported for the probability of loans elasticities (elementary schools): all were found relatively large but remain questionable due to non linearity and their approximation from a Taylor series expansion.

Source: Listings L11DP02, L11DPY4, L11DPYF, L11DPN1, L11DPOE, L11DPF1, L11DP8Y
L11DPS4, L11DPGY, L11DP00, L11DPXU, L11DPMG.

Policy Variables²

1. *The size and quality of the stock of printed materials* is measured by the number of books and periodicals per pupil, new acquisitions as a percentage of the total number of books, and the previous year's weedings as a percentage of the total number of books. These last measures are used as *proxies* for the quality of books; the hypothesis employed is that library personnel acquire books and dispose of others to improve the stock. Eliminating outdated or dilapidated books should increase the access to the remaining books. The size and quality of the stock are expected to influence directly and positively the probability that a library makes loans and the number of loans made.

2. *The library personnel* are classified as professionals, technicians or support staff working full-time or part-time. Professionals include professional librarians, educational advisors or specialists in learning materials and methods as well as teachers participating in their school library's operation. Technicians are employees who have received some training in library management and who work under the supervision of a professional. Library employees are also expected to influence directly and positively the output variables. Furthermore, as additional personnel should improve the efficiency of the management of the stock of books, interaction variables between the quantity and quality of book variables and the library personnel variables should also have a positive effect on the output variables. To illustrate the concept of interaction effect, let us consider the interaction variable defined as "weedings x full-time professionals." The use of this variable implies that the influence of the variable "weedings" on the probability of loans and on the number of loans also depends on the number of professionals working full-time in the school library.³

3. *The library organization and facilities* are taken into account by including the number of operating hours per week, the total library area, and the number of chairs and carrels available per pupil. The first two variables should influence the number of loans positively but the last two variables may simply induce pupils to work or read on the premises rather than borrow the materials for use outside the library.

A stronger negative effect is expected if the school encourages the formation of separate collections of books in classrooms, in addition to the existence of a central library. Finally, as for the previous class of factors, some interaction variables between library organization and facilities and (1) library personnel, (2) number of books, are included in the causal model.

Standardizing Variables

Users' characteristics are considered as playing an important role in determining the use of school libraries.⁴ Four characteristics of students are selected as particularly relevant:

1. *The age* of students is expected to influence loans positively, since older students are expected to read more both for classwork and leisure. However some empirical studies (Houle, 1980) on reading habits suggest the opposite hypothesis. The influence of age is captured through variables measuring the fraction of students in the lower grades at each level. These variables are expected to have a negative impact on loans.
2. *The socioeconomic variables* used are a socioeconomic status index attached to the area in which the school is located and an urbanization index for the same

area. The socioeconomic status ranges from strongly underprivileged to privileged. We expect that the more privileged the group is, the higher the probability of loans and the number of loans per pupil are. An index of urbanization is also part of the model to characterize the students' environment. Although the influence of the level of urbanization on school library loans is expected to be significant, the sign of this influence is not clear a priori.

3. Students who have chosen to enter *professional programs* in the latter years of their secondary education are expected to use school libraries less than students in general programs, *ceteris paribus*.
4. *The proximity of a public municipal library* could have opposing effects on the use of school libraries by pupils. The first effect is one of substitution, students borrowing fewer books from school libraries because they use the public library. The second is that the habit of frequenting the public library could increase the students' demand for reading materials.

The school characteristics selected as relevant here can be divided into three categories:

1. *School size*, measured by the number of teachers, is expected to have a significant influence on lending policies.
2. *Whether or not a school is public or private, English or French*, or belongs to an *integrated school board* (regrouping both elementary and secondary schools), is expected to influence the lending policies of the school library; however, we do not have a priori anticipations on the direction of this influence.
3. "*Polyvalent*" secondary schools offer a wider range of options in their curriculum. This may influence the reading behavior of students and also the lending policies of the library.

One category of variables we had hoped to include in our causal model of library use was that of pedagogical variables. It seems clear that teachers have a strong influence on students' reading and therefore on school library loans. Information on teachers' pedagogical methods and on their ability to use the library materials efficiently is, unfortunately, not available. Nevertheless, as it stands, the model includes important explanatory factors and policy instruments and is particularly suitable for empirical analysis as will be seen in the following sections.

The Data and Construction of Variables

The data used in this study came from two main sources: (1) a yearly school library survey yielding quantitative information on loans of printed materials, human resources, and physical facilities, and (2) a school data file providing information on the number of teachers and other school characteristics. For the academic year 1978-1979, this information was available for 877 public and private elementary school libraries and 350 public and private secondary school libraries, distributed across the province of Québec.⁵

To complement these data, the Canadian 1976 population Census and the Québec Municipality Library Repertory were used, from which were extracted area populations and information on the presence of municipal public libraries within reasonable distance of the schools. Furthermore, a socioeconomic index for the area in which each school in the sample is located⁶ was constructed. Table 1 summarizes the information for the key variables of the model.

An interesting point is that most (80.8%) of the school libraries included in our sample make loans to students at the secondary level whereas less than half of them

(47.5%) do so at the elementary level. However for those libraries making loans, the number of annual loans per pupil decreases substantially from 38.0 at the elementary level to 17.0 at the secondary level. The average number of books and periodicals is larger at the secondary level but the variations in per pupil annual loans among school libraries are also higher at this level.

The quality of the stock of books, measured by the number of acquisitions and weedings in the previous year as a percentage of the total number of books, varies very little across levels of schooling or between subsamples; one exception is the somewhat surprisingly high rate (15%) of new acquisitions for the full sample of secondary school libraries. The figures on library personnel clearly show that this component is a very scarce resource indeed.

An average school library opens for substantially longer hours per week and provides more space and a larger number of carrels at the secondary than at the elementary level.

The three socioeconomic strata are fairly well represented in the sample but the four chosen levels of urbanization are not. School libraries in large urban areas seem under-represented. Finally, as expected, both larger schools (based on teacher number) and a higher fraction of private schools are found at the secondary level.

Empirical Results

In discussing the results of the estimation of the model previously presented, a distinction is made between factors affecting the probability of a school library to make loans and the factors affecting the number of loans per pupil made over a school year. The statistical techniques used were probit analysis,⁷ to estimate the probability of loans equation, and ordinary least squares, to estimate the number of loans equation. Results are discussed according to the same classification of variables as in the section describing the explanatory model.⁸

The Quantity and Quality of Printed Materials

The number of books per pupil is one of the most important determinants of the number of loans made by school libraries at both the primary and secondary levels. Furthermore, although there is some departure from uniformity in the results, the positive effect of the number of books is much stronger when library personnel is present. For instance, the interaction between the number of books and (i) the technicians working full-time at the elementary level or (ii) the professionals working part-time at the secondary level, increases the positive impact of the stock of printed materials on library use.

This indirect effect of personnel is confirmed by the results on the quality of books at the elementary level. Unless interacting with personnel variables, the quality of books (% of new acquisitions and culling) does not have a strong positive significant influence on the number of loans. A puzzling result is the lack of significant positive coefficients for quality and personnel interaction variables (and even one significant negative coefficient) at the secondary level.

As for the interaction between the number of books and physical facilities, it exerts a positive effect in the case of books and the library area and a negative effect in the case of books and number of chairs at the elementary level; at the secondary level, there is a positive effect in the case of books and number of carrels. It is not surprising that a larger area, facilitating a better presentation of books, should contribute to increase the circulation of additional books.

Turning to the determinants of the probability that a school library makes loans, we find that the number of books or periodicals has no direct (positive) effect in elementary or secondary schools. The number of books does exert a positive influence, however, through the interaction with personnel. The interaction between books and library area, at the elementary level, and between books and number of carrels at the secondary level also increases the probability that some loans are made. At the elementary level, one quality dimension, new acquisitions, has a positive direct effect which is reinforced by mostly positive effects for the interaction variables between acquisitions and personnel. At the secondary level, the other quality dimension, weeding, has a negative direct effect which is more than compensated for by a positive indirect effect (through the interaction variables with personnel).

The elasticities in Table 2 illustrate clearly the relative importance of the quantity and quality of printed materials in explaining (i) whether or not a library makes loans, and (ii) the number of loans per pupil made in a year. Elasticities give the change in percentage terms of the variables explained in the model following a percentage variation in any of the nondichotomized explanatory variables.⁹ Use of the elasticities reveals that increasing the per capita number of books by 20% would lead to an increase of 10% in the number of loans per capita made by elementary and secondary school libraries. This is a strong reaction. The elasticity of the probability of loans with respect to weeding also appears to be quite high. However, as noted at the bottom of Table 2, approximate standard errors for elasticity estimates, in the case of the probability of loans, are large and therefore the results should be interpreted with some caution.

Library Personnel

An unexpected result of this study is the absence of any positive direct effect of personnel on loans at the primary level and the quasi-absence of such an effect at the secondary level. We have already discussed the generally positive indirect influence of personnel through the interaction with the size and quality of the stock of books. Such a positive indirect effect on the number of loans is also indicated by the positive coefficients of the interaction variable between the number of operating hours and personnel (part-time professionals and support staff) at the secondary level. On the other hand for other types of personnel (full-time professionals and part-time technicians), this interaction variable has a negative influence. Total elasticities of Table 2 summarize the information on the direct and indirect effects of library personnel on loans. The elasticities are rather small, which indicates that an increase in personnel would not lead to a substantial increase in loans, and that, moreover, some elasticities are negative.

Library Physical Facilities and Organization

The results on the effect on loans of the number of operating hours per week are as ambiguous as the results on personnel are inconclusive. The number of operating hours is a key element in explaining both the probability that a library makes loans and the number of loans made. The regression coefficients are positive and significant, especially at the elementary level, and the total elasticities of loans with respect to this variable are large. For instance, lengthening the hours of operation by 20% (going from 32.2 to 38.6 hours per week on average) would increase the number of loans by 17% in secondary schools.

Another organizational variable which has a significant influence is the presence of separate collections (outside the library). It reduced significantly the probability that an elementary school library makes loans. The physical facilities variables play a somewhat lesser role. The total library area per pupil increases the probability of loans at the elementary level and the number of loans at the secondary level. The number of chairs has a positive influence on library book loans while the number of carrels has the opposite influence at both schooling levels. A tentative explanation is that the presence of carrels may be an inducement to substitute in-library reading for borrowing.

User Characteristics

The influence of socioeconomic status is significant but it is not uniform across schooling levels. Results show, at the elementary level, that the higher the socioeconomic index, the higher the probability of loans and the number of loans. On the contrary, at the secondary level, the higher the socioeconomic index, the lower the probability of loans and the number of loans.¹⁰ The population size of the town or city where an elementary school is located influences positively the probability that the library makes loans but, for libraries already making loans, it reduces the number of loans per capita. The reverse is true at the secondary level although the pattern is somewhat less clear.

A higher proportion of students in the lower grades in the elementary and secondary schools increases the probability of loans. This unexpected result raises the question of the more or less intensive use of library materials by teachers at different grade levels.

The existence of a nearby municipal library influences negatively the probability of loans at both the elementary level and the number of loans at the secondary level. This might indicate that the actual relationship between public and school libraries is one of substitution rather than a complementary one for the younger clientele.

School Characteristics

The larger size of the school, measured by the number of teachers, increases the probability of a library making loans but does not significantly influence the number of loans. Libraries of private secondary schools, most likely because of their specific vocational nature, make fewer loans. The fact that a secondary school is polyvalent (that is, it offers a wider range of options) has a negative effect on the number of loans.

Conclusion and Policy Recommendations

This study illustrates the need to gather more and better data on variables associated with school library services. But even with incomplete information, the present study has shown that specific factors influence the use of the loan services offered by the school libraries.

The following conclusions and recommendations which policy makers should consider to complement their own personal judgements are worthy of note:

1. Among all variables, the number of books and the number of school library operating hours are the most important factors affecting positively the probability of loans and the number of loans made by school libraries.

2. Whether some types of library personnel are required to handle more books or have longer hours of operation is a technical question not addressed in the present study.¹¹ Although some positive indirect influence of personnel was indicated by the results, no direct influence of library personnel on the library output considered here was perceived.
3. For the physical facilities variables two results emerge: at the secondary level, the size of the library area has a positive influence on loans while, at the elementary level, the seating accommodation as represented by the number of chairs does so.
4. Some user characteristics and school characteristics influence the level of library services and this factor should be taken into account in designing policies.

In formulating policies to improve the economic efficiency of public and private school library policies, two further dimensions would have to be explored. First, one would want to know the influence of pedagogical methods and teachers' sensitization to the usefulness of school libraries on library use. Second, once a cost is assigned to each factor, a cost-benefit analysis should be performed so that public resources are allocated to library inputs efficiently. On that last matter, the nexus of number of schools, library personnel and costs must be seriously considered in light of the results obtained in this study.

Notes

1. Scattered references can be found in specialized library journals; for example, the *School Librarian* and the *School Media Quarterly*.
2. Some of those policy variables have been empirically considered in the public library literature; see, for example, Campbell and Shlechter (1979), Hayes (1979), and Powell (1978).
3. The reasoning is the same for "full-time professionals" whose influence can be reinforced by the number of "weeding." Formally, let Y : the number of loans per pupil, X_1 : the number of books disposed of, X_2 : the number of "full-time professionals." Assuming $Y = b_1X_1 + b_2X_2$ and with $b_1 = b'_1 + b''_1X_2$, we obtain finally $Y = b'_1X_1 + b''_1X_2X_1 + b_2X_2$, where X_2X_1 is the interaction variable. Now clearly this model cannot empirically be distinguished from the following: $Y = b_1X_1 + b_2X_2$ with $b_2 = b'_2 + b''_2X_1$ leading to $Y = b_1X_1 + b'_2X_2 + b''_2X_2X_1$.
4. This is also well documented in the public library literature; see, in particular, Stratton (1976) and D'Elia (1980a, 1980b).
5. Unfortunately, the city of Montreal, the largest urban community of the province of Québec, is under-represented in this sample.
6. Details and reference sources are given in Houle, Benyahia, and Montmarquette (1980), pp. 82-83.
7. Probit analysis deals with the problem of limited dependent variable as in the case when use of school library is defined equal to one and nonuse to zero in a regression framework; see Theil (1971, pp. 630-631) for details. D'Elia (1980b) has treated this problem in his analysis of use versus nonuse of public library services.
8. Complete regression results are available from the authors on request.
9. All the elasticities are calculated at the mean values of variables except for the elasticities of the probability of loans of secondary school libraries which are computed at the mean value of variables minus two standard errors (or set close to zero when a negative value results) to avoid being caught in a very flat portion of the probability curve, since most secondary school libraries do provide loans to their students.
10. Given the results, it is impossible to infer the existence of a substitution or complementary relationship between school library materials and reading materials available to the student at home.
11. On this particular problem, see Renner and Clark (1979).

This is a revised and condensed version of a report prepared for the Service général des moyens d'enseignement, ministère de l'Éducation du Québec, on *la mesure de l'impact des déterminants dans l'utilisation des services offerts par les bibliothèques scolaires*.

Financial assistance of the Centre de recherche en développement économique of the Université de Montréal is acknowledged for this revised version along with the technical assistance of Lina Di Pietro and Lisa Di Pietro.

References

- Campbell, D. E., & Shlechter, T. M. (1979). Library design influences on user behavior and satisfaction. *The Library Quarterly*, 47, 26-41.
- D'Elia, G. (1980a). A procedure for identification and surveying potential users of public libraries. *Library Research*, 2, 239-281.
- D'Elia, G. (1980b). The development and testing of a conceptual model of public library user behavior. *The Library Quarterly*, 50, 410-430.
- Hayes, R. M. (1979). The management of library resources: The balance between capital and staff in providing services. *Library Research*, 1, 119-142.
- Houle, R. (1980). *Rapport d'enquête sur les habitudes de lecture des élèves du secondaire*. Québec: Collection SREP, Gouvernement du Québec.
- Houle, R., Benyahia, H., & Montmarquette, C. (1980). *La mesure de l'impact des déterminants dans l'utilisation des services offerts par les bibliothèques scolaires*. Montréal: Centre de recherche en développement économique, Université de Montréal.
- Powell, R. R. (1978). An investigation of the relationship between quantifiable reference service variables and reference performance in public libraries. *Public Libraries*, 48, 1-19.
- Renner, C., & Clark, B. M. (1979). Professional and nonprofessional staffing patterns in departmental libraries. *Library Research*, 1, 153-170.
- Stratton, P. (1976). *Public libraries: Their structure, use and cost of service provision*. De Kalb, IL: Northern Illinois University.
- Theil, H. (1971). *Principles of econometrics*. New York: John Wiley & Sons.

NORMAN ROBINSON

Simon Fraser University

and

CAROLINE STACEY

Delta School District

Serving on the School Board: A Political Apprenticeship for Higher Office

The following question was investigated in this study: How do successful politicians who began their political careers as school board members and who have gone on to higher political office evaluate their school board experience in terms of its value as a political apprenticeship? Data were gathered from structured interviews with ten federal and provincial politicians who had once served as school board members; they agreed unanimously that school board experience was a valuable political apprenticeship for higher office. This was true regardless of the individual's political socialization or political recruitment experiences prior to school board service. The politicians felt that they had acquired valuable political skills and that their school board experiences had a strong politicizing effect on them. During their period of service on the board they had experiences which changed their political ambitions from office-bound ambitions (at the school board level) to upward-bound ambitions (at the federal or provincial level).

Researchers have given little attention to investigating the value of the school board experience as a political apprenticeship for those individuals who leave the school board and go on to higher political office. The reason for this neglect probably lies in the fact that the great majority of school board members do not have political ambitions above the school board level. For example, Ziegler, Jennings, and Peak (1974) found that only 7% of incumbent school board members had upward-bound political aspirations. In an earlier study which investigated the rea-

Norman Robinson is a professor of educational administration at Simon Fraser University. His major research interest is the politics of educational governance.

Caroline Stacey is a recent graduate of the Administrative Leadership Program at Simon Fraser University. She is currently working as a school administrator in Delta School District in British Columbia.

sons why people run for school board, McCarty (1959) found that less than 10% of school board candidates listed future political advancement as a motive for running. In a replication of the McCarty study done in British Columbia, McDonald (1972) found a similarly small percentage of candidates running to further their political careers.

The question of how many school board members actually do go on to higher political office has not been thoroughly researched. Existing evidence suggests that probably more school board members in Canada move on to higher political office than is the case in the United States. For example, Schlesinger (1966) found that few candidates for governorship and the U.S. Senate had a background of local government experience. In Canada, on the other hand, studies (Kornberg, 1967; Kornberg & Mishler, 1976; Kornberg, Smith, & Clarke, 1979) done on the backgrounds of members of the House of Commons reveal considerable experience in local government.

School board experience, as distinct from other forms of local government experience, has not been isolated in either the American or Canadian studies. As a result, no complete and accurate information exists on how many former school board members have gone on to higher political office in the United States or Canada. An inspection of the most recent Canadian parliamentary guide (Normandin, 1982) does reveal, however, that 23 of the present 282 members of the House of Commons are former school board members. About ten percent of the members of the provincial legislatures are former board members. Ontario, for example, has 17 former school trustees in its 125 member house.

Little also is known about school board members who seek higher political office but are unsuccessful in their efforts. No complete and accurate information on this question has yet been assembled. A recent study by Stacey (1983) did, however, document the number of school board members in British Columbia who offered themselves for election in the Canadian federal election in 1979 and in the provincial elections in 1979 and 1983. In the federal election, a total of six former school board members sought election to B.C.'s 28 seats in the House of Commons. Three of these former school board members were successful and became members of parliament. In the provincial election of 1979, there were 16 former school trustees seeking election to the 57 seats in the B.C. legislature; four were successful in winning seats. In the 1983 B.C. provincial election, 13 former trustees sought election and four were successful.

What becomes clear is that a small but important number of former school board members are successful in their quest for higher political office, some of whom later rise to fill high level political offices.

The Research Question

The research question that this study attempted to investigate can be stated as follows: How do successful politicians who began their political careers as school board members, and who have gone on to higher political office at the federal or provincial level, evaluate their school board experience in terms of its value as a political apprenticeship?

In studying this question, an attempt was also made to examine the political antecedents of the school board experiences of the politicians studied. That is, an examination was made of the political socialization experiences and the political recruitment experiences leading up to the politicians' school board experiences.

Methods

The Sample

The population for this study consisted of all former school board members, living in B.C., who are serving or have served in the House of Commons or the Legislative Assembly of British Columbia. In spite of a rather thorough search of the Canadian parliamentary guides over the past five years, it was not possible to identify the exact size of this population and to come up with an appropriate sampling procedure. As a result, a decision was made to interview all of those who could be contacted and who were willing to be interviewed. This turned out to be ten people.

In research terms, this sample of ten politicians is very much an accidental sample—it is not a random sample. It does possess, however, great diversity in terms of political characteristics which reflects the reality of political life in B.C. The sample had the following characteristics:

1. it contained representatives from all four political parties currently active in B.C.—four New Democrats, two Liberals, two Progressive Conservatives, and two Social Crediters;
2. five of the politicians had service in the House of Commons—four in the B.C. Legislature and one in both;
3. three of the politicians had service in either a federal or provincial cabinet;
4. three of the politicians had retired from politics; seven were still active;
5. four of the politicians had served only on the government side, three only on the opposition side, and three on both sides of the house;
6. there were eight men and two women; and
7. the period of service in political office at the federal or provincial level ranged from a low of two years to a high of 26 years.

Interviews

Structured in-depth interviews were used in this study. The interview schedule included questions regarding the political socialization, political recruitment and political apprenticeship dimensions of this study. The principal questions were as follows:

1. What were the important factors in the early *political socialization* of these politicians?
2. What were the important factors in the *political recruitment* process that led up to their school board service?
3. In what ways can service on a school board be considered a *political apprenticeship* for higher public office?

Results

Early Political Socialization

In political science research, political socialization is usually viewed as the process by which an individual acquires political values, attitudes, and interests. In addition, the individual acquires a knowledge of the political community with its institutions and actors. From childhood onward political socialization is a continuous process with the individual being placed under constant potential change through the influence of family, peers, school and other sources.

It was important in this study, first, to determine what factors played a significant part in the early political socialization of the ten politicians studied and, secondly, to assess the effect that these socialization experiences had on the initiation and development of the political careers of these ten individuals.

Findings

Family experiences played an important part in the early political socialization of eight of the ten politicians. Only two of the politicians came from families where there was no active interest in politics. For the other eight, an interest in politics was characteristic of family life. Typical of the comments describing the family situation were the following:

My parents were both very active politically. They talked politics at the dinner hour . . . so I grew up with it.

My father and mother were both actively involved in politics. My mother mainly in a supportive role, my father mainly in the union, as president of his local. He was deeply involved in union politics, and as I said, ran for alderman in the city.

Four of the politicians came from family situations where a parent or close relative had held public office.

I came from a very political family. My father was a Tory and he served on the school board too. The old time Tories used to meet at our house with my father and I used to love to hear them talk.

My father was on the school board for several years. He was chairman. Later he was elected to the House of Commons.

The interviews with the politicians revealed as well that the majority of them (six of the ten) had become politically active in a formal party well before their entry into school board affairs—a few in adolescence and others, after high school, usually in university. Typical of the experiences described are these three comments:

My first public experience was when I was in grade eight and I was asked to give the toast to the graduating class. From then on I became active in student council affairs in high school. I joined the C.C.F. at a very young age . . . so at the age of fourteen or fifteen I was attending political meetings.

I had been involved in the Liberal Party since university days and followed through afterwards . . . I remained an active member and was president of a very small local at the time I became a school trustee.

My father ran for alderman a number of times . . . and I delivered campaign literature for him. I joined the party after that. So I've always been involved in politics in some way, at university and elsewhere. At the time I ran for trustee I was president of the local N.D.P. club.

Discussion

The findings of this study reveal quite clearly that the majority of the politicians interviewed came from politically active families. In most cases, politics was part of family life and in nearly half of the families, either a parent or close relative had held public office.

These findings support those of other studies which show that the level of political interest and activity is much higher in the families of future politicians than is the case among members of the general public. A study by Kornberg and Mishler (1976) of Canada's 28th Parliament (1972-1974) showed that nearly 66% of the

members had fathers with a strong interest in politics and 38% had mothers with a similar high level of interest. Furthermore, Clark, Price, and Krause (1976), in a study of provincial legislators in all ten Canadian provinces, found that 35% had parents or other relatives who had held public elective office.

What the findings of this study do not support is the conclusion of Zeigler, Jennings, and Peak (1974) that school board members do not come from politicized family backgrounds. Their study of 490 school board members in 83 American school districts showed that only one-third of the board members came from families where the father was "very much interested" in politics and only one-fifth of board members put their mothers in that category.

What is quite clear from the findings of this study is that the overwhelming majority of these future school board members underwent early political socialization experiences that were strong and pervasive. What is not clear is whether most individuals who seek school board membership have similar kinds of political socialization experiences. Are strong and pervasive early political socialization experiences found only in the backgrounds of board members whose political ambitions become upward-bound (i.e., oriented to attaining higher political office) or are they found as well in the backgrounds of board members whose political ambitions remain office-bound (i.e., restricted to the school board)? As well, it is unclear whether there are differences between Canada and the United States in terms of the political socialization experiences that prospective office seekers undergo prior to standing for election to a school board.

Political Recruitment to the School Board

The fact that an individual comes from a highly politicized family environment and has strong and pervasive political experiences in youth and early adulthood does not mean that an individual will embark on a political career. In any society there are thousands of individuals who come from highly politicized backgrounds but only a few ever stand for public office. Thus, Milbrath (1965) and Bowman and Boynton (1966) suggest that future political actors have to cross a political activity threshold. They have to be stimulated to stand for public office either through the encouragement of others or through the arousal of their own inner feelings. This crossing of the political activity threshold constitutes the process of political recruitment. In this study, the focus was on what factors were important in the political recruitment process that lead up to service on the school board by these politicians.

Findings

Election to the school board constituted the beginning of a political career for nine of the ten politicians. Only one had prior experience in a public elective office.

Prior to their election to school board, three of the politicians were nonactive in community affairs; the remaining seven had a varied pattern of community activity. In their study of school board members, Zeigler, Jennings, and Peak (1974) examined the community activities of members prior to their election to office. These researchers classified the community activities into three categories: (1) civic/business activities (i.e., involvement in service clubs or business groups); (2) political/governmental activities (i.e., political party involvement or service on public boards); and (3) educational activities (i.e., involvement in educational affairs as a teacher, administrator, or member of a parent-teacher group). Quite clearly these are not dis-

crete categories and it is not possible to categorize the community activities of the individuals in this study into three tidy areas. It was possible, however, to identify from the interviews what the primary focus of their community activities was. For the seven who were active in community activities, two can be classified as active in civic/business affairs and five in political/governmental affairs. None could be classified as being active in education in spite of the fact that three were former school teachers. In all three of these cases, the individuals involved had a long history of political involvement and party membership before they entered teaching.

A number of writers (Zeigler, Jennings, & Peak, 1974; Goldhammer, 1964) have pointed out that a disproportionally high number of school board members come from "education" families, that is, families where parents or relatives are employed in educational positions. This is certainly true of this study. As has been noted already, three members had themselves been former teachers, one had been an educational administrator, and four others had close relatives in educational positions. This meant that only two of the ten politicians had no personal or family involvement in education.

In terms of their recruitment to stand for election to school board, the politicians interviewed were asked to name persons or groups who encouraged their candidacies. Six had been approached and encouraged to run, the other four indicated that their candidacies were self-initiated. Of the six who had been approached, three were approached by informal groups (parents, friends, neighbors), two were approached by political leaders (mayors in both cases), and one was approached by a formal group (ratepayers' association).

Issue differences with the incumbent school board played an important role in stimulating the school board candidacies of six of the ten politicians interviewed. One whose candidacy was self-initiated said:

I was always very vocal, it was part of my background. And one day I was sitting in the staff room complaining about the school board and somebody said, "You should get on that school board and then maybe you would stop complaining so much about it." So after I left teaching I decided to run for trustee. I became a trustee because there were things going on in education in our district that I didn't like.

Four of the ten politicians who decided to run for school board did not have issue differences with the incumbent board members. Their reasons were often highly personal. One said:

After the (federal) election, six of us sat around and conducted a post mortem on why we hadn't been able to get anyone elected. I suggested that we should start seeking municipal office . . . so that we would have a track record that would indicate that we could hack public life. . . . So I decided to run for school board and eventually I became an M.P. . . . I planned a career in politics, beginning with the school trusteeship.

Kornberg and Mishler (1976) did an analysis of the motives of candidates who sought office and were successful in becoming members of parliament in the Canadian federal election of 1972. Although there was much overlap in categories, they found that the candidates were motivated by the following (in order of importance): (1) "good government" considerations; (2) ideological considerations; (3) "challenge of office" considerations; (4) party pressure; and (5) pressure from others.

In this study the six candidates who were approached to run for school board tended to have different motives for running than the four individuals whose candidacies were self-initiated. The motives of those approached could be roughly categorized as follows: (1) pressure from others (three candidates); (2) "good govern-

ment” considerations (two candidates) and (3) ideological considerations (one candidate). For the four self-initiated candidates, the breakdown on motives is as follows: (1) “challenge of office” considerations (three candidates); and (2) ideological considerations (one candidate).

Discussion

The pattern of political recruitment to school board service for the individuals in this study differs considerably from the pattern of recruitment for school board members reported elsewhere (Goldhammer, 1964; Jennings & Zeigler, 1971; Cistone, 1974; Zeigler, Jennings, & Peak, 1974). The two most extensive studies of the political recruitment of school board members are those of Cistone (1974), done in Ontario, and Zeigler, Jennings, and Peak (1974), done in the U.S.A. on a nationwide basis.

An exact comparison cannot be made between these three studies because each study did not ask the same questions in all areas. The limited comparisons that can be done, however, do reveal certain important differences between the results of this study and the results of the earlier studies of Cistone and Zeigler, Jennings, and Peak. In terms of political recruitment to the school board, the results of this study show a higher level of political/governmental activity in the community prior to school board service, more self-initiated candidacies, no encouragement to run from incumbent school board members or professional educators, and more issue differences with incumbent board members.

The only major similarity in the results of this study and the results of the Cistone and the Zeigler, Jennings, and Peak studies is that all three show that school board members are drawn predominantly from education families. Over 80% of board members in all three studies had relatives in education and a sizeable minority of board members had been professional educators.

One finding of this study that deserves particular comment relates to the total absence of any activity on the part of incumbent school board members to approach or encourage potential candidates to run for office. A number of writers (Crain, 1968; Goldhammer, 1964; Jennings & Zeigler, 1971; Tucker & Zeigler, 1980) have noted that there is a strong tendency on the part of school boards to perpetuate themselves. This is done by encouraging the candidacies of carefully selected individuals. In this study, none of the politicians interviewed indicated that they received any support or encouragement from incumbent school board members to run for office.

School Board Service as a Political Apprenticeship

In this study an attempt was made to determine whether school board experience served as a political apprenticeship for the successful federal and provincial politicians studied, and if so, in what ways the experience was valuable to them.

Findings

All ten politicians were of the strong opinion that their school board experience was an invaluable political apprenticeship. For all but one of the politicians, service on the school board was their first experience in public office and for them it was a powerful and mind-expanding experience. As one of the politicians described it:

I look back on my school board years as one of the most interesting and stimulating times of my life. It was my first entrance into politics and everything was to my mind so excit-

ing. I was almost overwhelmed with the fact that I had been given the opportunity to sit on the board and to make decisions.

The politicians mentioned particular skills they acquired as a result of their school board experience. Foremost among these were skills in dealing with people, particularly in the areas of coalition building and conflict resolution. One of the politicians put it this way:

You may have the best ideas in the world as a school trustee but until you get half of the people on your side, it won't go anywhere. So you have to work with people whom you may not even care to talk with, with whom you have nothing in common, but with whom you have to work with in order to accomplish what you want. As a politician, learning that when I was a trustee was very useful.

In describing the human relational skills they acquired as a result of their school board experience, other politicians said:

You learn to work in groups, how to come to decision making, how to share strong ideas . . . I think it's excellent political experience.

You become adept at . . . juxtaposing conflicting roles and dealing with people who may or may not have some opposing view.

Several of the politicians interviewed made specific reference to the skills they acquired in dealing with professional experts. One politician said:

I learned not to have blind faith in expertise. A lot of teachers, administrators and experts didn't know what they were talking about. I learned a built-in suspicion to the expert as opposed to the common sense approach of people.

Another politician added,

I learned to be skeptical with experts and with conclusive findings.

All of the politicians expressed the view that they acquired valuable skills in financial matters. Typical of the comments relating to the acquisition of financial skills is the following:

I learned to be resourceful. How to husband monies and how to zero in on the source of a problem and how to then go about solving it.

The political skills of coalition building, conflict resolution, assessment of expert evidence, and financial management were identified by the politicians as the most valuable skills they acquired as a result of their school board experience. Certain other important political skills that were also acquired but which were mentioned less frequently include skills in chairing meetings, communicating to groups, listening to others.

Another important finding of this study was that school board experience had a politicizing effect on the individuals involved. They found that they liked politics, they could be successful at it, and this interest and success encouraged them to consider running for higher political office. One of the politicians put it succinctly by saying:

People either become political animals or they don't after serving on a board . . . I became one in the sense that I enjoyed politics.

Another politician described how success in the school board experience can serve as a stimulus for the seeking of higher political office:

If you are successful as a trustee you are motivated to go on. . . . If you are successful in improving society in whatever area, then it's only natural that if you see you can contribute in some larger way you will be motivated to continue.

With the exception of one individual, none of the politicians consciously planned to use the school board as a stepping stone to higher political office. The stirring of political ambition came about as a result of the school board experience. One politician described the process as follows:

Most of the people I know that have come from school boards and (city) councils didn't enter the council or trustee job thinking they would run for higher office. I think they sought the job and that was the end to the goal. As they got into it, they enjoyed it and perhaps became popular in their own area, be it in their community or their political party, and saw the opportunity present itself one day to go on. I don't think that people who go out for school board and council really see it as a means to an end. I think they go out to change the world, and are either in the end defeated, drop out, or go on. It's like any other process, there is room to keep on climbing. Any person who is ambitious, when they start on a job, will be aiming for a better job.

Success and a sense of accomplishment in the school board member role was thus of great importance in encouraging these individuals to seek higher public office. Three of the individuals interviewed did not feel, however, that they had accomplished as much as they had wanted. They felt that this was less a function of themselves and more a function of the limited power school boards have. One of the former school board members put it this way:

My feeling was that when I went in and became a school trustee most of the decisions had already been made by the time they were presented to us at the school board level. The government makes decisions on financing and there is very limited room to move as a trustee. I think somebody once said that fifteen percent of the budget is yours to do whatever you want with, but only if you raise taxes. I don't think that it is a particularly satisfying job.

An examination was also made of the effect that electoral defeat had on the political ambitions of these former school board members. Only two of the ten politicians had ever suffered electoral defeat, and only one had been defeated in a school board election. The effects of defeat tend to be temporary and no real antidote to political ambition. The two politicians in this study who had experienced defeat described their reaction as follows:

Well, I tangled with some teachers and was defeated the second term. That was it, I had had enough. Later, however, I was asked to run for city council and eventually became mayor.

Yes, I lost two elections. I learned a lot by losing. When you understand why you lost, you can do something about it. . . . It's depressing, but not devastating, to lose. You have to understand that most people run to win. You are perhaps depressed for a week or so, but my God, it's only an election!

Finally, an attempt was made to determine whether school board experience made the individuals in this study desirable to political parties as potential federal or provincial candidates. This was a difficult question to sort out as many of the individuals were engaged simultaneously in school board affairs and party affairs at the constituency level. Several of the politicians did indicate, however, that their school board experience was very important in terms of their desirability as a candidate for higher office. One said:

If I hadn't been an elected official for ten years I'm sure that members of my party would not have approved me to run. A party wants to win and when they are looking for someone, they are looking for someone who is electable.

Another described his experience in these terms:

I think the fact that I was a school trustee influenced people to ask me to seek further office. I had always been involved in politics, long before I was a school trustee but basic-

ally at the party organization level rather than in elected office. Getting elected as a school trustee introduced me to a whole new realm of politics. Then when the former candidate decided not to run again, Tommy Douglas approached me, and the local party president approached me, and I said "sure I'll make a go for it."

Discussion

To this time, researchers have given little attention to school board service as a political experience. What the results of this preliminary study show is that it can be a powerful experience in terms of the acquisition of political attitudes and skills. Regardless of the nature of their earlier political socialization or political recruitment experiences, the individuals in this study were unanimous in their agreement on the value of the school board experience as a political apprenticeship. They became more politicized as a result of the experience, and in addition, acquired skills which they deemed to be invaluable in their future political careers.

Summary and Implications

This study examined the value of the school board experience as a political apprenticeship for a group of ten politicians who went on to successful careers at the federal or provincial level. These politicians agreed unanimously that school board service was a valuable political experience for them. It was one in which they acquired needed political skills and attitudes. They also acquired a certain desirability as potential candidates for higher public office.

Only one of the ten politicians in this study deliberately set out to build a political career using the school board as a stepping stone. The remainder did not have upward-bound political ambitions at the time they became board members—they simply wanted to be school trustees. Yet during the course of their tenure as board members they had experiences which changed their political ambitions from office-bound ambitions to upward-bound ambitions. This was true regardless of the nature of their prior political socialization or political recruitment experiences.

In this study we examined the school board experiences of individuals who have gone on to higher political office. We also examined their political socialization and political recruitment backgrounds. Most school board members do not go on to higher office. Quite a few try, but most fail. For everyone who serves on a school board, the experience is, among other things, a political experience. For all the individuals in this study, the experience was useful and for some it was also powerful and mind-expanding. For those board members who do not go on to higher office, or for those who fail in their efforts to do so, the nature of school board service as a political experience may well be different. It may be that their political socialization and political recruitment experiences differ markedly from those of the politicians in this study. Only further research can clarify these issues and our future research efforts will focus on these issues.

References

- Bowman, L., & Boynton, G. R. (1966). Recruitment patterns among local party officials. *American Political Science Review*, 60, 667-676.
- Cistone, P. J. (1974, October). School board member recruitment: The case of Ontario. *Journal of Educational Administration*, 12(2), 43-56.

- Clarke, H. D., Price, R. G., & Krause, R. (1976). *The role socialization of the Canadian political elite: A note on provincial legislators*. Unpublished paper, University of Windsor.
- Crain, R. L. (1968). *The politics of school desegregation*. Chicago: Aldine.
- Goldhammer, K. (1964). *The school board*. New York: Center for Applied Research in Education.
- Jennings, M. K., & Zeigler, H. D. (1971). Response styles and politics: The case of the school boards. *Midwest Journal of Political Science*, 15, 290-321.
- Kornberg, A. (1967). *Canadian legislative behavior*. New York: Holt, Rinehart & Winston.
- Kornberg, A., & Mishler, W. (1976). *Influence in Parliament: Canada*. Durham, NC: Duke University Press.
- Kornberg, A., Smith, J., & Clarke, H. D. (1979). *Citizen politicians: Canadian party officials in a democratic society*. Durham, NC: Carolina Academic Press.
- McCarty, D. J. (1959, September). School board membership: Why do citizens serve? *Administrator's Notebook*, 8(1).
- McDonald, K. R. (1972). *The interrelationships of motives for seeking school board membership, mode of school board operation and community type*. Unpublished master's project, Simon Fraser University.
- Milbrath, L. (1965). *Political participation*. Chicago: Rand McNally.
- Normandin, A. L. (Ed.). (1982). *Canadian Parliamentary guide*. Ottawa: Editing and Administration, P.O. Box 3453, Station C.
- Schlesinger, J. (1966). *Ambition and politics: Political careers in the United States*. Chicago: Rand McNally.
- Stacey, C. (1983). *The school trusteeship as a political experience*. Unpublished master's project, Simon Fraser University.
- Tucker, H. J., & Zeigler, H. D. (1980). *The politics of educational governance: An overview*. Eugene, OR: ERIC Clearinghouse on Educational Management.
- Zeigler, L. H., Jennings, M. K., & Peak, G. W. (1974). *Governing American schools*. North Scituate, MA: Duxbury Press.

FRANK M. MERRITT

and

R. STEVE McCALLUM

University of Southern Mississippi

The Relationship Between Simultaneous-Successive Processing and Academic Achievement

The present study investigated the relationship between simultaneous-successive information processing and academic achievement among 157 college students. A factor analysis of 6 tasks comprising the simultaneous-successive battery yielded a two-factor solution. Four processing groups were formed using factor scores as criteria. Cumulative grade point average (GPA) and composite American College Testing (ACT) scores represented the dependent (achievement) variables. Results indicate that level of simultaneous and successive processing is related systematically to GPA; high levels of both processing modes are necessary for high GPA achievement. On the other hand, simultaneous processing seems relatively more important for high ACT performance. These results are discussed as they relate to task demands of the ACT and GPA.

The study of human learning often progresses by using scientific models consistent with the existing Zeitgeist. For example, several current cognitive models utilize computer programming to simulate learning. Similarly, increasingly sophisticated neurophysiological research has made neuropsychological models more relevant. Recently, these two types of models have been integrated, creating a neuropsychological model of human learning which is typically described in computer processing terminology. Founded upon A. R. Luria's clinical observations (Luria, 1966a, 1966b), this model has been elaborated and extended by J. P. Das

Frank M. Merritt is currently at the University of Southern Mississippi Psychology Department completing doctoral work in Clinical Psychology. He maintains active research interests in neuropsychologically-based learning models.

R. Steve McCallum is currently an Associate Professor of Psychology at the University of Southern Mississippi and Associate Editor of the *Journal of Psychoeducational Assessment*. He also maintains active research interests in neuropsychologically-based learning models.

and colleagues via empirical research (see Das, Kirby, & Jarman, 1979). This model is often termed the Luria-Das Information Processing Model.

The structural and processing features of the Luria-Das model can be briefly described as follows. The brain is composed of three units. Unit 1 is primarily composed of upper/lower brain stem organs and is responsible for the arousal or motivational function; Unit 2 is located in the posterior region of the neocortex including the occipital, temporal, and parietal lobes. Unit 2 structures receive information, analyze it into elementary components, and code it into a symbolic/physiological form useful to the entire brain. Information is analyzed and coded within Unit 2 in either a simultaneous or a successive fashion. Simultaneous integration refers to the synthesis of separate elements into groups, often with spatial characteristics. Any part of the gestalt and its relative position is immediately surveyable. On the other hand, successive integration refers to the processing of information in a serial order so that information is surveyable only in a temporal, orderly manner, with knowledge of each bit of information dependent upon the previous bit. Finally, Unit 3 is comprised of the prefrontal lobes, and is responsible for planning, executing, and evaluating problem-solving strategies.

Conceptualizing human learning within this comprehensive model should provide guidance in the selection, validation, and interpretation of measurement instruments. Contrasted to traditional empirical approaches to intellectual assessment which are primarily based on predictive validity, the Luria-Das model utilizes relevant theoretical and empirical research to conceptualize cognitive functioning. Therefore, results from testing are amenable to interpretation within a meaningful conceptual framework and subsequent remedial strategies are easily established (e.g., Kaufman & Kaufman, 1979; Krywaniuk & Das, 1976).

Obviously, empirically derived tests which predict achievement are valuable to educators. Tests yielding good predictive validity *and* information relevant for interpreting performance would be even more valuable. Studies exploring the relationship between simultaneous-successive processing described within the Luria-Das model and academic achievement offer evidence that such tests may be possible in the near future. Results from elementary school age children indicate that simultaneous-successive processing and achievement are related, but in a complex way. For example, when a simultaneous-successive battery is administered along with reading achievement tests, results reveal that successive processing is the mode often employed by beginning readers; more sophisticated readers seem to rely on both modes (Das, Kirby, & Jarman, 1979).

The present study was designed to investigate the relationship between simultaneous-successive processing and academic achievement in college students. Questions focused on whether simultaneous-successive processing ability varied systematically with achievement. Individuals demonstrating high simultaneous and successive processing skills were expected to earn the highest achievement scores whereas those demonstrating low simultaneous and low successive processing skills were expected to earn the lowest achievement scores; consistent with the non-hierarchical nature of the model (Das et al., 1979; Jarman, 1978), individuals demonstrating high simultaneous-low successive ability should perform similarly to those exhibiting low simultaneous-high successive ability.

Method

Subjects

Participants included 157 juniors and seniors enrolled in Educational Psychology courses at a mid-sized southeastern university. This sample was composed of 53 males, 104 females; 133 were white, 24 black. Ninety-seven were education majors. Data from males and females were combined for analyses (see Merritt & McCallum, 1983; Kirby, cited in Das et al., 1979, for justification).

Simultaneous and Successive Tasks

Six tasks commonly used to assess simultaneous-successive processing were administered in group format. The tasks included:

Raven Progressive Matrices. This test is described as a culture-reduced measure of nonverbal reasoning (Raven, 1958). Participants indicated which of a given set of alternatives correctly completes a visual pattern. Fifteen of the 60 available items (every fourth item) were chosen for presentation.

Memory for Designs. The nine items from the Bender Visual Motor Gestalt Test (Bender, 1946) were presented for four seconds each. Following the presentation, examinees constructed the figures from memory.

Figure Copying. The last 12 figures from the Developmental Test of Visual Motor Integration (Beery & Buktenica, 1967) were presented for 30 seconds each and copied during that time.

Free Recall. Ten lists containing six monosyllabic words were presented orally. All 60 words appeared at the 1st, 2nd, or 3rd grade levels of the Slosson Oral Reading Test (Slosson, 1963) or the Basic Sight-Word Test (Dolch, 1942). The order of recall was not critical.

Serial Recall. Ten additional lists of 6 monosyllabic words from the 1st, 2nd, and 3rd grade levels of the Slosson Oral Reading Test and the Basic Sight-Word Test were presented orally. The order of recall was critical.

Digit Span-Forward. This task required recall of increasingly longer lists of digits. The 14 lists ranged from 4 to 10 digits in length. Each was recalled and written immediately after oral presentation.

Rationale for Task Selection

Of the six tasks, Raven Matrices, Figure Copying, and Memory for Designs are considered indices of simultaneous processing. Serial Recall, Free Recall, and Digit Span-Forward are purported to be measures of successive processing. Evidence for construct validity of the battery comes from several sources. Luria's clinical observations of brain-damaged individuals clearly indicate the existence of these two modes of processing, and the six tasks described above exhibit face validity; i.e., they appear to assess the two basic processing modes described by Luria. And though the simultaneous-successive tasks have been psychometrically classified into parallel categories such as nonverbal-verbal, visual-auditory, and reasoning-memory, the two factors which have emerged from analyses can be appropriately termed simultaneous and successive processing. For example, simultaneous processing is utilized in verbal tasks such as the perception of lexical ambiguity (Das & Jarman, 1980), in auditory tasks such as syllogisms (Cummins, cited in Das, Kirby, & Jarman, 1975), and in memory tasks such as memory for designs (McCallum & Merritt, 1983). Additional convergent and discriminant validity is provided by Das et al. (1979).

Achievement Measures

Achievement measures were cumulative grade point average (GPA) and the American College Testing (ACT) composite score. The ACT composite score reflects assessment of English and Math usage, and Social Studies and Natural Science Reading skills; it is typically administered to high school seniors and valued for predictive validity of college success. The scores were obtained from student records kept by the university.

Procedure

Before the simultaneous-successive battery was administered, a questionnaire requesting demographic data was completed by each student. The six tasks were then administered in group form, with simultaneous tasks alternating with successive tasks. The order of presentation of the tasks shifted systematically with each presentation. Administration was standardized using a slide projector and taped audio, and required approximately 55 minutes.

Task scores were subjected to a principal components analysis; factors with eigenvalues greater than 1.0 were rotated according to a varimax criterion. Grouping of subjects was accomplished on the basis of mean splits of factor scores. Therefore, group assignments reflect relative ability to process information. Performance of the group members were compared on ACT scores and cumulative GPA, the dependent variables. Contrasts predicted from the model employed one-tailed tests of significance.

Results

As in a previous study employing a college student sample (McCallum & Merritt, 1983), the factor analysis of task scores yielded two factors labelled simultaneous processing (defined by Raven Matrices, Memory for Designs, and Figure Copying) and successive processing (defined by Free Recall, Serial Recall, and Digit Span-Forward). (See Table 1 for factor loadings, means, and standard deviations.) Factor scores were subsequently obtained and became the criteria for assignment to one of four groups—high simultaneous-high successive (HiSim-HiSuc), high simultaneous-low successive (HiSim-LoSuc), low simultaneous-high successive (LoSim-HiSuc), and low simultaneous-low successive (LoSim-LoSuc). Because group assignments were based on mean splits of factor scores, groups reflect *relative* ability to process information.

As previously mentioned, in order for performance to be consistent with the model, members of the HiSim-HiSuc group should have earned significantly higher achievement scores than members of any other group. Conversely, members of the LoSim-LoSuc group should have earned scores significantly lower than those from any other group. These ten contrasts were predicted by the model and allowed use of a priori *t* tests.¹ Performance of the group members within the two intermediate or “off-diagonal” groups should have been similar because the model suggests a nonhierarchical arrangement of the two processing modes. That is, the two groups high on only one processing dimension should have performed about equally. These two contrasts employed two-tailed criteria. Means, standard deviations, and *t* ratios for all group comparisons are presented in Table 2.

Results employing the cumulative GPA were consistent with predictions from the model for all contrasts. Table 2 presents the relevant values for the cumulative GPA analyses.

TABLE 1
MEANS, STANDARD DEVIATIONS, AND FACTOR LOADINGS
FROM SIX COGNITIVE TASKS

Tasks	<u>M</u>	<u>SD</u>	Successive	Simultaneous
Digit Span	8.18	2.21	.85	-.02
Free Recall	49.87	5.54	.76	.34
Serial Recall	43.57	7.43	.81	.24
Raven Matrices	8.08	2.10	.11	.77
Memory for Designs	28.14	1.51	.22	.58
Figure Copying	7.24	2.24	.10	.85

TABLE 2
COMPARISON OF THE MEANS OF CUMULATIVE GPA AND COMPOSITE ACT
FOR INFORMATION PROCESSING GROUPS

GROUPS	ACT				GPA			
	<u>M</u>	<u>SD</u>	<u>t</u>	<u>df</u>	<u>M</u>	<u>SD</u>	<u>t</u>	<u>df</u>
HiSim-HiSuc	18.92	3.80	.07	63	3.04	.48	2.47**	82
HiSim-LoSuc	18.85	4.58			2.77	.50		
HiSim-HiSuc	18.92	3.80	3.35**	60	3.04	.48	2.21*	77
LoSim-HiSuc	15.70	3.43			2.79	.49		
HiSim-HiSuc	18.92	3.80	4.73**	63	3.04	.48	4.37**	90
LoSim-LoSuc	14.38	3.77			2.56	.56		
HiSim-LoSuc	18.85	4.58	2.69**	47	2.77	.50	.15	63
LoSim-HiSuc	15.70	3.43			2.79	.49		
HiSim-LoSuc	18.85	4.58	3.83**	50	2.77	.50	1.71*	76
LoSim-LoSuc	14.38	3.77			2.56	.56		
LoSim-HiSuc	15.70	3.43	1.27	47	2.79	.49	1.79*	71
LoSim-LoSuc	14.38	3.77			2.56	.56		

* $p < .05$

** $p < .01$

Results from analyses using the composite ACT scores were more ambiguous. Examination of the relevant values in Table 2 reveal only partial support for the processing model. Three predictions were supported; three were not. The mean composite ACT for the HiSim-HiSuc group was significantly higher than the mean scores obtained by the LoSim-HiSuc group and the LoSim-LoSuc group. Also, the HiSim-LoSuc group mean score was significantly higher than the mean obtained by the LoSim-LoSuc group. However, contrary to prediction, the mean HiSim-HiSuc

group score was not significantly greater than the mean from the HiSim-LoSuc group. Nor was the mean from the LoSim-HiSuc group significantly higher than the mean from the LoSim-LoSuc group. Finally, the two intermediate group means were not equivalent; the HiSim-LoSuc group mean was significantly higher than the mean from the LoSim-HiSuc group.

Discussion

The strong two-factor solution obtained from the initial factor analysis of the six cognitive tasks was anticipated. The solution is consistent with a rational analysis of task demands and with findings from numerous studies (see Das et al., 1979). Since these factors did emerge, further analyses were conducted to investigate the relationship of college achievement to simultaneous-successive processing.

When cumulative GPA was employed as the dependent measure, all predictions were confirmed. The cumulative GPA reflects performance in a wide variety of skill areas; apparently both simultaneous and successive processing substantially mediate broadly-based classroom performance. Individuals within the HiHi group performed better than those high in only one area; those high in only one processing dimension out-performed those in the LoLo group; and those in the two "off-diagonal" groups performed equally well. These results provide support for the model and are consistent with the contention that the modes are nonhierarchical.

Three of the six predictions using the composite ACT score were not supported—the HiSim-HiSuc group did not earn a higher mean ACT score than the HiSim-LoSuc group; the LoSim-HiSuc group did not earn a higher mean ACT score than the LoSim-LoSuc group; and, finally, the HiSim-LoSuc group did earn a significantly higher mean score than the LoSim-HiSuc group. Thus, the contrasts not conforming to predictions substantiate the relative importance of the simultaneous processing mode for attaining higher ACT scores. But because test reviewers (e.g., Wallace, 1975) have criticized the ACT for a heavy reading component, the exceptions may be interpreted within the context of previous results which emphasize the importance of simultaneous processing skills for complex reading (Cummins & Das, 1977). That is, if the ACT is interpreted as a measure of sophisticated reading skills, then the superordinate status of simultaneous processing in the ACT measure may be explained by the theory.

In summary, the present results are consistent with the Luria-Das model. Specifically, there appears to be a relationship between cognitive processing ability and academic achievement of college students. In general, higher simultaneous processing is important for superior ACT performance whereas higher simultaneous and successive processing are necessary for superior GPA.

Implications

Rigorous experimental hypotheses testing of the Luria-Das model is clearly needed. Thus far, educators have been relatively unsuccessful in identifying appropriate treatment x aptitude models (Reynolds, 1981). But research into cognitive functioning within this model appears promising. For example, once simultaneous or successive deficits have been identified, remedial strategies become apparent. Means of identifying simultaneous-successive deficits include the use of the battery and methodology described within the present paper. A second method, currently becoming available, is called the Kaufman Assessment Battery for Children (K-ABC) (Kaufman & Kaufman, 1983). The availability of this instrument and the in-

creasingly abundant research findings investigating the model should hasten efforts to explore the efficacy of the model within educational and clinical settings.

Notes

1. The model-based hypotheses of the present study were based on the assumption that the dependent achievement measures required both processing modes in about equal proportions. Therefore, the use of multiple a priori *t*'s was justified, and the concomitant increased probability of Type I error was tolerated. In retrospect, this assumption was not justified for the ACT dependent measure (i.e., simultaneous processing is apparently more important), but the *t* ratios of the ACT measure exceed the more conservative $\alpha < .01$ value in all cases of significance.

References

- Beery, K. E., & Buktenica, N. A. (1967). *Developmental test of visual motor integration: Student test booklet*. Chicago: Follett.
- Bender, L. (1946). *Visual motor gestalt test*. New York: American Orthopsychiatric Association.
- Cummins, J., & Das, J. P. (1977). Cognitive processing and reading difficulties: A framework for research. *Alberta Journal of Educational Research*, 23, 245-256.
- Das, J. P., & Jarman, R. F. (1980). Coding and planning processes. In M. Friedman, J. P. Das, & N. O'Connor (Eds.), *Intelligence and learning*. New York: Plenum Press.
- Das, J. P., Kirby, J., & Jarman, R. F. (1975). Simultaneous and successive syntheses: An alternative model for cognitive abilities. *Psychological Bulletin*, 82(1), 87-103.
- Das, J. P., Kirby, J. R., & Jarman, R. F. (1979). *Simultaneous and successive cognitive processing*. New York: Academic Press.
- Dolch, E. C. (1942). *Basic sight-word test*. Champaign, IL: Garrard.
- Jarman, R. F. (1978). Level I and Level II abilities: Some theoretical reinterpretations. *British Journal of Psychology*, 69, 257-269.
- Kaufman, A. S., & Kaufman, N. L. (1983). *Kaufman assessment battery for children*. Circle Pines, MN: American Guidance Service.
- Kaufman, P., & Kaufman, D. (1979). Strategy training and remedial techniques. *Journal of Learning Disabilities*, 12, 63-66.
- Krywaniuk, L. W., & Das, J. P. (1976). Cognitive strategies in native children: Analysis and intervention. *Alberta Journal of Educational Research*, 22, 271-280.
- Luria, A. R. (1966a). *Higher cortical functions in man*. New York: Basic.
- Luria, A. R. (1966b). *Human brain and psychological processes*. New York: Harper & Row.
- McCallum, R. S., & Merritt, F. M. (1983). Simultaneous-successive processing among college students. *Journal of Psychoeducational Assessment*, 1, 85-93.
- Merritt, F. M., & McCallum, R. S. (1983). Sex-related differences in simultaneous-successive information processing. *Clinical Neuropsychology*, 5, 117-119.
- Raven, J. C. (1958). *Standard progressive matrices: Sets A, B, C, D, and E*. London: H. K. Lewis.
- Reynolds, C. R. (1981). Neuropsychological assessment and the habilitation of learning: Consideration in the search for the aptitude x treatment interaction. *School Psychology Review*, 10, 343-349.
- Slosson, R. L. (1963). *Slosson oral reading test*. New York: Slosson Educational Publications.
- Wallace, W. L. (1975). Review of the ACT test battery of the American College Testing Program. In O. K. Buros (Ed.), *Intelligence tests and reviews*. Highland Park, NJ: Gryphon Press.

ESSAY REVIEWS

LOGO AND THE DEVELOPMENT OF GENERAL PROBLEM-SOLVING SKILLS

Linda R. Krasnor

and

John O. Mitterer

Brock University

How to structure children's encounters with computers and what to expect from those encounters are questions with far-reaching educational implications. Particularly high expectations have been aroused by LOGO, a graphics-oriented structured language which is imbedded in an educational philosophy stressing play, discovery and mastery (Papert, 1980; Groen, 1978). LOGO allows children to direct a "turtle" around a video screen to create designs. Once a small set of primitives has been learned, children can "teach" the computer to draw designs by defining procedures. Each child can develop a set of procedures which can be combined, just like the primitives, into higher order procedures to make complex designs. Thus the child gradually builds a personal computer language while learning fundamental programming concepts.

Papert (1980) has argued that when children experience LOGO they learn much more than a computer language. The LOGO experience was designed explicitly to facilitate the learning of powerful ideas, skills and heuristics which transcend the immediate task environment and can be applied in other problem-solving situations. This is an exciting possibility since curriculum designers have traditionally focused on specific content rather than general problem-solving skills (Brown, Bransford, Ferrara, & Campione, 1983; Larkin, 1980).

It is not implausible that some skills learned in programming may transfer to other types of problem-solving since it has often been noted that computer programming is a type of problem-solving. The steps in designing a program closely

Linda Rose Krasnor is a Developmental Psychologist with research interests in social interaction during problem solving activities. She is at Brock University.

John Mitterer is a Cognitive Psychologist interested in cognitive development. Recently, he has become interested in the applications of cognitive and developmental theories to education.

parallel generally recognized steps in the problem-solving process (Statz, 1973; Chait, 1978; Papert, 1980). In Polya's (1957) classic model of problem-solving, four basic steps were outlined: understanding the problem, devising a solution, carrying out the solution plan, and evaluating the solution. Similarly, the programmer must understand the programming task and be able to represent the problem. The program must be devised, entered into the computer, and debugged.

Several skills have consistently been advanced as powerful ideas which can be acquired through the LOGO experience (Papert, 1980). One is the strategy of breaking problems into manageable subproblems. Another is means-end analysis or the systematic planning of actions to achieve goals (e.g., Sacerdoti, 1977). A third is debugging, the principle that problem solutions can be successively refined. The development of a more positive attitude to errors is a fourth powerful idea. Programming bugs are recognized as being unavoidable and useful in providing information. Such a reconceptualization of failure implies that the child's sense of mastery and motivation would not be negatively affected by errors.

In this paper, we will provide a framework for assessing the extent to which these skills learned through LOGO may be transferred to other problem-solving contexts. Focusing on the turtle geometry aspects of LOGO, we will examine some of the principles governing the generalization of learning as they apply to the LOGO experience and will consider the existing empirical evidence for the transfer of these skills to other domains. We will also consider how transfer may depend on individual differences among children. Finally, we will isolate specific components of the complex LOGO experience in an attempt to specify which components may be responsible for transfer if and when it does occur.

The Generalization of Learning

As Simon (1980) has made clear, the notion of general knowledge is really the notion that some kinds of knowledge can be transferred from the context of acquisition to a wide variety of other situations. Thus, a consideration of the circumstances which will facilitate the transfer of learning from one situation to another will in effect be a consideration of the factors governing the generality of knowledge. From this perspective, recent research has most often failed to produce evidence for the transfer of knowledge between domains (see Brown, et al., 1983). It now appears that much domain specific, nontransferable knowledge is required for expertise in any domain. As Simon (1980) has suggested, "No one, no matter how intelligent, skilled in problem solving, or talented, becomes a chess grandmaster without 10 years of intense exposure to the task environment of chess" (p. 82).

While it is possible that the basic strategy of training general skills for transfer may not be as effective as previously supposed, there are, nevertheless, many factors which have long been recognized to influence transfer (Ellis, 1965). One important principle is that transfer between tasks requires that some of the processes or knowledge used in the tasks be essentially identical (Thorndike & Woodworth, 1901). A second is that the learner must recognize that a new problem situation is similar to one encountered previously (e.g., Woodworth & Schlosberg, 1954). It has commonly been found that little spontaneous transfer occurs between well-formed problem isomorphs (e.g., Reed, Ernst, & Banerji, 1974; Hayes & Simon, 1977). If explicit instruction is given in the range of applicability of a general principle (Brown, 1978; Brown & Campione, 1981), however, or if learners are encouraged to state a general rule (Gagne & Smith, 1962; Thorndyke & Stasz, 1980), then transfer may be

improved. Third, the degree of transfer may depend on the completeness of original learning (Brown et al., in press). Finally, the degree of transfer may depend on exposure to a variety of situations in which the specific skill is useful. Training in multiple contexts has been found to improve transfer, when compared to relatively uniform learning contexts (Belmont, Butterfield, & Borkowski, 1978; Brown, 1974, 1978; Brown, Campione, & Day, 1981).

Therefore, transfer (if it occurs at all) would be expected between highly similar domains under circumstances where learners were fluent with the skills in question, had practised those skills in several situations, and were explicitly aware of the general utility of the skills. With these considerations in mind, to what extent would the four powerful ideas from LOGO be expected to transfer to other problem domains?

The only condition for transfer to be consistently met by such an experience is that LOGO problems are similar to problems in other domains. However, this similarity may well be at a relatively abstract level and many students will need explicit guidance in order to recognize correspondences. LOGO instructional experiences typically have not included any attempt to explicitly teach students in the use of these skills. Rather, tutors rely on subtle hints and structuring of problems to achieve this end. This may mitigate against transfer rather than facilitating it.

LOGO instruction is generally short-term; a typical LOGO experience consists of 20 to 40 hours of hands-on experience. There is thus little opportunity to ensure sufficient mastery of programming fundamentals to successfully apply means-end and sub-goal analysis to many different problems. Also, it is quite possible that the original learning has taken place in a very narrow context, especially for those children who work only with turtle graphics. The lack of practice across a variety of situations may reduce the opportunity for transfer even further.

The situation is somewhat improved with respect to the powerful ideas of debugging and a positive attitude towards bugs. In LOGO instruction, tutors are advised to encourage exploration after errors and to minimize the negative consequences of errors. Children are exposed to debugging very early in the instructional sequence, since they make errors as soon as they begin. Students thus have the opportunity to edit procedures repeatedly. Considerable expertise with these skills should be built up. Lack of explicit instruction in their general applicability and a lack of practice in different contexts may reduce the degree of transfer, however, even for these powerful ideas.

On balance, we simply cannot predict how much transfer of these powerful ideas to expect from the LOGO experience. However, we can observe that some of the most important factors in promoting the transfer of other skills are not present in the LOGO experience as it is most commonly described in the literature. Clearly, there are many research opportunities presented by the questions raised here. As the following review of the literature will reveal, few researchers have examined these important issues.

Evidence for the Generality of LOGO Learning

To date the literature concerning LOGO consists primarily of testimonials (e.g., Goldenberg, 1980; Watt, 1982), curriculum guides (e.g., Birch, 1980), manuals (e.g., Abelson, 1982) and a few studies containing considerable anecdotal evidence. Although researchers have explored the use of LOGO with adults (Brown & Rubenstein, 1974; Lukas & Feurzeig, 1972; Austin, 1976; duBoulay, 1977), children

have received the most attention. Most of the available evidence refers to LOGO learning rather than the transfer of LOGO learning to other contexts. This body of work has provided anecdotal evidence that the powerful ideas mentioned above can be learned by children and a suggestion that these skills can be transferred to other domains.

The work of the M.I.T. LOGO group has been typical of this research approach. These researchers have focused on case studies (e.g., Lawler, 1980) and the creation of curriculum materials (e.g., Bamberger, 1979) rather than experimental findings. The Brookline project is a representative example. This collaboration between the M.I.T. LOGO group and the public schools of Brookline, Massachusetts, resulted in detailed anecdotal descriptions of the learning of 16 children without systematic experimentation or statistical analysis (see Papert, Watt, diSessa, & Weir, 1979).

A few quotes from the interim report of this project (Papert, Abelson, Bamberger, diSessa, Weir, Watt, Hein, & Dunning, 1978) will suffice to illustrate these descriptive data. The comments concern a single child, Ray. The first two were made by one of the authors (Watt) who noted,

From the comments made by his teachers, and from observations made by myself and others in the LOGO classes, I have a model of Ray as a boy who has chosen to react to academic difficulties by adopting a pose of indifference, and refusing to take responsibility for his work or behavior in school. (Appendix 1, Individual profiles; Ray, p. 1)

Watt goes on, "to *speculate* that from his earliest school experiences, Ray has been afraid of failure, and especially of *appearing* to fail" (p. 2) (author's italics). Finally, after learning LOGO, one of Ray's teachers commented, "The breakthrough for him in LOGO, the success he has had, is powerful information for me . . . he has produced the best piece of writing I've seen from him" (p. 12). Such descriptions definitely suggest the transfer of important abilities from LOGO to another domain, but very little objective analytic evidence is provided.

This research approach is consistent with both published (Papert, 1980) and private (Solomon, 1982, personal communication) claims that traditional experimental research methodologies are inappropriate for documenting the effects of LOGO. Much of the other published work on LOGO suffers from a similar lack of empirical documentation (e.g., Feurzeig & Lukas, 1972; Weir & Emanuel, 1976). This view has several sources: (a) that traditional psychological and educational testing methods are not appropriate for measuring LOGO-facilitated changes; (b) that group-based designs and statistical analyses will obscure the wide individual differences in children's learning patterns and adult teaching styles; and (c) that standardization of LOGO teaching procedures will substantially detract from the flexibility and individualization of the LOGO experience.

Only a handful of studies have produced empirical evidence relevant to the claim that learning LOGO produces general effects. Of these, the Edinburgh LOGO project (e.g., Howe, O'Shea, & Plane, 1979) focused on the development of mathematical skills and has produced mixed evidence of benefits while the thesis work of Statz (1973) and Chait (1978) focused on the development of general problem-solving skills.

Of the two reports, that of Statz (1973; see also Statz, Folk, & Seidman, 1973) more closely approximated a classical experimental design. Problem-solving was investigated by administering pre- and posttests to a group of 16 grade 4 children who had participated in LOGO sessions and to a similar group who remained in the normal class environment. The two groups were administered four problem-solving

tests. The Twenty Questions Game (Bruner, Rose, Oliver, & Greenfield, 1967) was used to investigate the child's ability to organize concepts hierarchically, as revealed through constraint-seeking questions. The Tower of Hanoi puzzle (e.g., Simon, 1975) was used to test the ability to develop a systematic solution strategy. To assess permutation abilities, the children were also asked to generate all possible orders of sets of numbers. Finally, a word task was used to assess classification skills. Statz (1973) predicted that learning LOGO would lead to improved performance on all of these problem-solving measures. While she found trends in the anticipated direction for all of the tasks, only the permutation and the word puzzle tasks reached significance.

Chait (1978) investigated comparable issues with five grade 7 children who were taught individually. These children were administered pre- and posttests but no control group was employed. The test battery included a test of basic arithmetic skills, an alternate uses test, and a set of geoboard problems including duplications, rotations, and mirror-imaging. The results revealed no significant changes in test scores.

The entire LOGO literature merits considerable criticism from an experimental point of view. The general lack of objective measurement makes the currently available work useful only as a first step in evaluating the effects of LOGO. The work of Statz reveals several problems which limit the usefulness of her work. First, the training procedure she used was modified partway through the study, making assessment difficult. Second, her measures of problem-solving were not tightly tied to any extant approach to problem-solving (e.g., Polya, 1957; Newell & Simon, 1972), although they might be considered to reflect the ability to carry out means-end analysis. The work of Chait is similarly open to criticism. A small sample size and the lack of a control group would have made the results difficult to interpret even had they reached significance. The test battery was also not tied to any theory of problem-solving. It is not clear why any of the tasks she chose should reveal any change as a result of learning LOGO.

To summarize, there is as yet no good evidence that any of the powerful ideas mentioned above generalize to other domains. This is either because existing research reports do not convincingly test for generalization (as in the case of means-end analysis) or because the relevant research has not yet been done. Clearly, in light of the potential impact of LOGO, these issues cannot remain unaddressed, and their resolution promises to be a fruitful area for future investigation.

Individual Differences and the Generality of LOGO Learning

One of the striking aspects of the current literature has been the extensive range of applicability claimed for LOGO. Learning LOGO is presumed to benefit normal children as young as three or four years of age. While there is considerable anecdotal evidence for the effectiveness of LOGO in teaching programming to such young children, more controlled research has indicated that preschoolers do not have the necessary spatial concepts to effectively use the turtle's intrinsic geometry and are unable to connect specific keyboard activity to robot movement (Gregg, 1978).

Children who are unable to utilize the intrinsic geometry system or to understand the relation between program commands and turtle movement would not achieve the mastery and control objectives of the LOGO experience. Children with inadequate language, spatial, and mathematical skills would have limited possibi-

ties for going beyond immediate programming modes. While it is likely that elementary computer literacy could be achieved, relatively little generalizable problem-solving skill would be expected without learning structured programming.

Young and relatively poor learners may need to be taught both specific *and* general skills in a given content domain in order to show transfer (Brown et al., 1983). More able learners may only need to be taught more specific skills, since they are more likely to possess general heuristics already. More advanced learners may also spontaneously prepare for transfer by doing such things as hypothesizing applications and using analogies. This suggests that younger children need to be explicitly taught general skills, as well as skills specific to the turtle world, before transfer could be reasonably expected. The discovery-oriented nature of the LOGO experience does not provide this explicitness for general skills. Thus, only older students may be able to apply LOGO skills to other domains.

Generalization of problem-solving skills may well also depend on the child's Piagetian developmental level. On one hand, Papert (1980) has argued that LOGO may offer to the child a way of concretizing the abstract, a very powerful tool for the attainment of formal operational thought. On the other hand, it is quite possible that only children who have reached the formal operational stage will be capable of fully utilizing LOGO (e.g., Folk, 1972). Both positions suggest that an optimal age for LOGO instruction may be the onset of the formal operational stage—about 12 years of age (Inhelder & Piaget, 1958). It is reasonable to assume that some general problem-solving heuristics will require formal operational thought. A child at a pre-operational level would be unlikely to achieve the hierarchical thinking necessary for structured programming. Such thinking may well be required to master subgoal analysis and mean-end planning. Even concrete operational children would likely find it difficult to manipulate symbols logically and to grasp concepts of recursion, multiple inputs, and permutation at an abstract level. Yet it is at this more advanced level that generalization of these skills would most likely occur. In contrast, it is likely that much younger children would be capable of understanding and benefiting from the powerful ideas of debugging and adaptive response to errors since it is less obvious how these skills may depend upon a foundation of formal operational thought. Although this is an issue of theoretical and applied import, to date there has been no systematic investigation of the relationship between developmental level and the generalization of problem-solving skills learned through LOGO.

In an early investigation, Folk (1972) has directly examined the possibility that the ability to learn LOGO fully in the first place may depend on the developmental level of the child. Unfortunately, Folk's unpublished thesis provides little more than a suggestive set of hypotheses. Thus, the extent of LOGO learning which may be expected for a child of a specific age or developmental level is still an open question. It is possible that only formal operational children will be able to learn LOGO fully and transfer LOGO ideas such as means-end analysis and subgoal analysis to other contexts. Adaptive response to failure, however, may be learned and generalized by much younger children.

In addition to a wide range of typical children, there has been a keen interest in the use of LOGO with special groups of children such as autistic, educable retarded, and dyslexic children (e.g., Goldenberg, 1980). The appeal of LOGO for these hard-to-treat groups is understandable. LOGO's novelty and immediate appeal, however, cannot ensure the sustained interest necessary to develop general problem-solving

skills. Without an understanding of the potential of LOGO with typical children, its use with atypical children may be built on relatively uncertain ground. It is likely that such special groups of atypical children would be much less likely to develop general problem-solving skills than normal children. Clearly, this issue provides much scope for future research.

Identifying the Components of LOGO Which Produce General Learning

A detailed task analysis of the LOGO experience will be required before the problem of generalization can be completely specified. As a first step, we have isolated several educationally potent components of the experience, each of which may be sufficient to induce cognitive and behavioral change. Two of these components are aspects of the LOGO language itself, two are aspects of the accompanying educational philosophy, and the final one is the motivating character of LOGO. The combination of these diverse factors in the LOGO experience makes it difficult to assess the unique contribution of LOGO *as a language* to learning and transfer.

The first feature of the language is the turtle, which is intended to provide the child with a way to "concretize the abstract" (Papert, 1980). By encouraging the child to "play turtle," syntonic learning is made possible. Associated with the turtle is intrinsic geometry or turtle geometry, a non-Cartesian approach to geometry (Abelson & diSessa, 1981). In turtle geometry, the state of the turtle is specified without respect to an external reference frame. This makes it easier to begin graphics programming since only the commands, LEFT, RIGHT, FORWARD, and BACK are required to specify turtle state, commands which derive meaning from the child's own physical experience. In addition, graphics-based programming environments may have a great appeal due to the potency of visual experience. The use of such a concrete medium may facilitate the ability to engage in means-end and subgoal analysis by making the task of completely conceptualizing a problem much simpler.

The second feature of LOGO is the ease with which structured programming is possible. New procedures can be created at will and freely combined with other procedures to create complex procedures more easily than in any other available language. In a compiled language like PASCAL, for example, a listing of the entire program (including all subprocedures) must be brought together for compilation. This is not necessary in LOGO, in which defined procedures can be treated just like primitives in developing more complex procedures. Subgoal and means-end analysis are clearly facilitated by such capabilities.

These two components of LOGO jointly create a programming environment characterized by immediate feedback and total control. As a LOGO graphics program is executed, each step of the program is displayed. Most bugs stand out and may be debugged right away. This should facilitate the acquisition of the idea that debugging is a useful problem-solving strategy. Compiled, text-oriented programming may not allow such efficient feedback. Such an environment, which is responsive to attempts at control, may allow children to develop a sense of mastery and competence (Bronson, 1974). In fact, it is one of the major claims of LOGO proponents (e.g., Papert, 1980) that the microworld of the turtle is one in which the child can feel powerful and successful.

In addition to these two components, several other elements of the LOGO experience can be identified which are not inherently tied to the language itself. One such component is a stress on play and exploration; the second is the small group nature of the LOGO experience.

In the LOGO environment, the child is free to experiment without risking any real-life consequences. This atmosphere of play and exploration is aided by allowing the child to select projects, thus creating a personal involvement and commitment which may nurture greater success. The child's affect is almost always described as positive, and the turtle world is certainly nonliteral. As such, the LOGO experience satisfies most major definitions of play (Krasnor & Pepler, 1980). The child is encouraged to try out new means of achieving goals and to combine old means to achieve new goals. This may result in a more flexible and adaptive mastery of LOGO than a more regimented approach, since it has been shown that free play with materials can lead to enhanced creativity and problem-solving (e.g., Dansky & Silverman, 1973, 1975; Smith & Dutton, 1979; Sylva, Bruner, & Genova, 1976). Thus it is likely that children who learn LOGO in an exploratory, play-based mode will be more adept at solving programming problems than children who are taught LOGO in a more directive form. In addition, adaptive attitudes toward failure and a willingness to debug problem solutions should be facilitated.

In most implementations of the LOGO curriculum, children share equipment and are encouraged to cooperate with peers to complete small group projects. This social arrangement can be considered to be another major component of the LOGO environment. Group problem-solving can be an intensive experience in itself, and may be an important factor in any changes which have previously been attributed to LOGO. The child has the opportunity to learn by observing the nature and outcome of others' actions (Bandura, 1977). In addition to benefiting from peer suggestions, the child may also engage in teaching others. This tutoring may help the child to consolidate and practise skills (Allen, 1976). Further, there is a need for the child to communicate clearly and effectively with peers, both in order to teach and to negotiate joint programming projects. Practice in these skills may lead to improvement in communication ability.

The negotiation and confrontation present in social interaction is thought to be an important factor in cognitive development, especially in the growth of decentration and perspective-taking skills (Piaget & Inhelder, 1969). The effects of small-group project-oriented interactions thus may be powerful ones, and there has been no control for its influence in prior research. Thus it is quite possible that children who learn LOGO in a small group will become more adept at programming in LOGO and at referential communication than children who learn LOGO alone. In addition, increased peer social status may result from attaining expertise in a new skill, and by the formation of new friendships through shared cooperative experience.

Finally, one of the assumed advantages of learning problem-solving skills through LOGO programming is that it is fun. This motivation may be a critical component of success in teaching LOGO and in producing any resulting transfer. A distinction should be made, however, between motivation for initially establishing contact with the LOGO experience, and the type of motivation which is needed for sustaining interest in LOGO over extended periods of time. There is little question that microcomputers and turtle graphics will provide sufficient novelty to motivate initial contact with the language. This initial motivation must carry the child through the period necessary to learn the basic fundamentals of the language and its technology. For a young child, this is not a trivial accomplishment. The child must learn the keyboard symbols, understand simple editing, recognize the need for exact syntax and supplying both direction and quantitative information. If the child does not continue efforts through this technique period, mastery of the skills neces-

sary to implement programming ideas and successfully solve problems will not develop.

To summarize, it is quite possible that the powerful ideas of debugging and positive response to failure can be transmitted simply by exposure to play-oriented, small-group experiences, even without any LOGO programming. However, the other powerful ideas of means-end analysis and subgoal analysis seem more tightly tied to the language itself. Even here, though, it is quite possible that noncomputer problem-solving curricula stressing these heuristics can be just as successful as LOGO training. It is in the motivational component, perhaps, that LOGO has a special advantage over other methods of teaching problem-solving. Alternately, it is possible that the overall instructional package offered by the LOGO experience, blending as it does so many potent instructional devices into one curriculum, provides by far the best context for facilitating the learning and transfer of these powerful ideas. Only the systematic investigation of the contribution of these and other components of the LOGO experience will result in definitive answers.

Conclusions

In our review of current research, we have found no evidence that the LOGO experience with turtle geometry will facilitate the development of general problem-solving skills. Further, there is some theoretical reason for doubting that such skills will be developed by the LOGO experience. Until further research resolves this question, educators interested in using LOGO in this way should tailor the LOGO environment to facilitate transfer. In general, the effective use of LOGO to develop problem-solving skills likely depends on three central conditions: (a) the use of techniques which foster awareness of the general utility of problem-solving heuristics; (b) selection of an appropriate group of children; and (c) the inclusion of the major components of the LOGO environment, since it is possible that LOGO programming per se will not be sufficient to produce transfer.

Tutors should explicitly emphasize how the skills learned in LOGO are relevant for problems outside the turtle world, how children may recognize appropriate similarities among problem types, and how LOGO skills might be adapted for other types of problems. Children should be encouraged to solve a variety of graphic and nongraphic problems, both in and out of the LOGO environment. Furthermore, it is likely that longer exposures to LOGO will increase mastery of general programming skills, which may enhance transfer. Extended training periods may be particularly critical for weak students, who will need more time to master the specific content of LOGO and to learn the general transfer skills. Much more research will be required to specify a curriculum using LOGO to promote general problem-solving skills.

On the basis of current evidence, it appears that only formal operational children of at least 12 years of age will be able to benefit fully from attempts to foster general problem-solving through LOGO. For younger or less able children, expectations for LOGO should be focused on mastery, adaptive reaction to failure, and computer literacy.

Attempts to fit LOGO into more traditional curricula (e.g., by grading projects, or predetermining the sequence of concepts learned) may reduce its potential benefit. If this is done, children may not have adequate opportunity to explore and play with LOGO, although play and exploration may be the most critical component for developing powerful ideas. Similarly, isolating the child with a single computer or otherwise regulating peer interaction in the interests of classroom management may limit the motivational power of LOGO and the benefits of group problem-solving.

Although we have focused upon LOGO's possibilities for enhancing general problem-solving and have attempted to highlight the potential limitations of LOGO for this purpose, there are many other reasons why children of a wide range of abilities and developmental levels should have an opportunity to experience LOGO. First, LOGO is fascinating in its own right. Further, LOGO may be a superior language for attaining computer literacy and introducing programming fundamentals, especially for younger children. Future research contrasting LOGO with other languages such as BASIC will be interesting. Finally, further research may show that the LOGO experience promotes the attainment of other educational objectives, such as fostering self-esteem, communication abilities, and a sense of mastery.

So far, there has been a tendency to overlook the need for objective and systematic examination of the LOGO experience and its effects. Without this appraisal, there is a distinct possibility that exciting learning opportunities may be lost. In the enthusiasm surrounding the availability of microcomputers and the development of languages such as LOGO, educators have envisioned dramatic and pervasive benefits for children. These expectations, if untempered, may eventually work against the productive use of computers in education. Unrealistic beliefs concerning the benefits of LOGO may result in frustration and disillusionment. Without analytic empirical investigations of LOGO, educators will lack the knowledge needed to design efficient and effective ways of using LOGO.

Our traditional knowledge of the principles of learning and transfer are still relevant, in spite of the futuristic nature of the educational technology. Past research, some of which dates to the beginning of this century, provides an important foundation for structuring LOGO learning environments and for predicting realistic outcomes. This blending of the old and the new provides an exciting and possibly unique research opportunity.

The authors especially wish to thank Joan Preston for her early participation on our LOGO project. In addition, Cynthia Solomon, Sharnee Chait (and LCSi), Allison Birch (and the Phoenix School), Sandra Wills, Dave Darvill, and Dale Burnett provided many stimulating thoughts. The preparation of this article was funded in part by Social Sciences and Humanities Research Council of Canada strategic seed grant #499-82-1010.

References

- Abelson, H. (1982). *LOGO programming for the Apple II*. Peterborough, NH: Byte/McGraw Hill.
- Abelson, H., & diSessa, A. (1981). *Turtle geometry: The computer as a medium for exploring mathematics*. Cambridge: M.I.T. Press.
- Allen, V. (1976). Children helping children: Psychological processes in tutoring. In J. Levin & V. Allen (Eds.), *Cognitive learning in children: Theories and strategies*. New York: Academic.
- Austin, H. (1976). Teaching teachers LOGO: The Lesley experiments. *M.I.T. LOGO Memo #23*.
- Bamberger, J. (1979). LOGO music projects: Experiments in musical perception and design. *M.I.T. LOGO Memo #52*.
- Bandura, A. (1977). *Social learning theory*. Englewood Cliffs, NJ: Prentice-Hall.
- Belmont, J., Butterfield, E., & Borkowski, J. (1978). Training retarded people to generalize memorization methods across memory tasks. In M. Gruneberg, P. Morris, & R. Sykes (Eds.), *Practical aspects of memory*. London: Academic Press.

- Birch, A. (1980). *LOGO curriculum for the Lincoln-Sudbury Schools, Mass.* Unpublished manuscript.
- Bronson, W. (1974). The growth of competence: Issues of conceptualization and measurement. In K. Connolly & J. Bruner (Eds.), *The growth of competence*. London: Academic Press.
- Brown, A. (1974). The role of strategic behavior in retarded memory. In N. R. Ellis (Ed.), *International review of research in mental retardation* (Vol. 7). New York: Academic Press.
- Brown, A. (1978). Knowing when, where, and how to remember: A problem of metacognition. In R. Glaser (Ed.), *Advances in instructional psychology* (Vol. 1). Hillsdale, NJ: Erlbaum.
- Brown, A., Bransford, J., Ferrara, R., & Campione, J. (1983). Learning, remembering and understanding. In J. Flavell & E. Markman (Eds.), *Carmichael's manual of child psychology* (Vol. 1). New York: Wiley.
- Brown, A., & Campione, J. (1981). Inducing flexible thinking: A problem of access. In M. Friedman, J. Das, & N. O'Connor (Eds.), *Intelligence and learning*. New York: Plenum.
- Brown, A., Campione, J., & Day, J. (1981). Learning to learn: On training students to learn from texts. *Educational Researcher*, 10, 14-21.
- Brown, J. S., & Rubenstein, R. (1974). *Recursive functional programming for the student in the humanities and social sciences* (Technical Report No. 27a). Department of Information and Computer Science, University of California.
- Bruner, J. (1972). The nature and uses of immaturity. *American Psychologist*, 27, 687-708.
- Bruner, J., Rose, R., Oliver, P., & Greenfield, P. (1967). *Studies in cognitive growth*. New York: Wiley.
- Chait, S. (1978). An analysis of children's problem solving in a graphics oriented computer programming environment. Unpublished master's thesis, McGill University, Montreal, Quebec.
- Dansky, J., & Silverman, I. (1973). Effects of play on associative fluency in preschool-aged children. *Developmental Psychology*, 9, 38-49.
- Dansky, J., & Silverman, I. (1975). Play: A general facilitator of associative fluency. *Developmental Psychology*, 11, 104.
- duBoulay, B. (1977, March). Learning teaching mathematics. *Mathematical Teaching*, 52-57.
- Ellis, H. (1965). *The transfer of learning*. New York: Macmillan.
- Feurzeig, W., & Lukas, G. (1972). LOGO: A programming language for teaching mathematics. *Educational Technology*, 12, 39-46.
- Folk, M. (1972). Influences of developmental level on a child's ability to learn concepts of computer programming. Unpublished doctoral dissertation, Syracuse University, Syracuse, NY.
- Gagne, R., & Smith, E. (1962). A study of the effects of verbalization in problem-solving. *Journal of Experimental Psychology*, 63, 12-18.
- Goldenberg, P. (1980). *Special technology for special children*. Cambridge, MA: M.I.T. Press.
- Gregg, L. (1978). Spatial concepts, spatial names, and the development of exocentric representations. In R. Siegler (Ed.), *Children's thinking: What develops?* Hillsdale, NJ: Erlbaum.
- Groen, G. (1978). The theoretical ideas of Piaget and educational practice. In P. Suppes (Ed.), *Impact of research in education: Some case studies*. Washington, DC: National Academy of Education.
- Hayes, J., & Simon, H. (1977). Psychological differences among problem isomorphs. In N. Castellan, J. D. Pisoni, & G. Potts (Eds.), *Cognitive theory* (Vol. 2). Hillsdale, NJ: Erlbaum.
- Howe, J., O'Shea, T., & Plane, F. (1979). Teaching mathematics through LOGO programming: An evaluation study. *DAI Research Paper No. 115*.
- Inhelder, B., & Piaget, J. (1958). *The growth of logical thinking from childhood to adolescence*. New York: Basic.

- Krasnor, L., & Pepler, D. (1980). The study of children's play: Some suggested future directions. In K. Rubin (Ed.), *Children's play: New directions for child development* (Vol. 9). San Francisco: Jossey-Bass.
- Larkin, J. H. (1980). Teaching problem solving in Physics: The psychological laboratory and the practical classroom. In D. T. Tuma & F. Reif (Eds.), *Problem solving and education: Issues in teaching and research*. Hillsdale, NJ: Erlbaum.
- Lawler, R. (1980). The progressive construction of mind (One child's learning: addition). *M.I.T. LOGO Memo #57*.
- Lukas, G., & Feurzeig, W. (1972, October). Computers and learning to learn. *Proceedings of the second City University of New York Conference, Innovations in Educational Technology*.
- Newell, A., & Simon, H. A. (1972). *Human problem solving*. Englewood Cliffs, NJ: Prentice-Hall.
- Papert, S. (1980). *Mindstorms: Children, computers and powerful ideas*. New York: Basic.
- Papert, S., Abelson, H., Bamberger, J., diSessa, A., Weir, S., Watt, D., Hein, G., & Dunning, S. (1978). Interim report of the LOGO project in the Brookline Public Schools. *M.I.T. LOGO Memo #49*.
- Papert, S., Watt, D., diSessa, A., & Weir, S. (1979). Final report of the Brookline LOGO project. Part II: Project summary and data analysis. *M.I.T. LOGO Memo #53*.
- Piaget, J., & Inhelder, B. (1969). *The psychology of the child*. New York: Basic.
- Polya, G. (1957). *How to solve it*. New York: Anchor.
- Reed, S., Ernst, C., & Banerji, R. (1974). The role of analogy in transfer between similar problem states. *Cognitive Psychology*, 6, 436-450.
- Sacerdoti, E. (1977). *A structure for plans and behavior*. New York: Elsevier North-Holland.
- Simon, H. (1975). The functional equivalence of problem-solving skills. *Cognitive Psychology*, 7, 268-288.
- Simon, H. (1980). Problem solving in education. In D. T. Tuma & F. Reif (Eds.), *Problem solving and education: Issues in teaching and research*. Hillsdale, NJ: Erlbaum.
- Smith, P., & Dutton, S. (1979). Play and training on direct and innovative problem solving. *Child Development*, 50, 830-836.
- Statz, J. (1973). *The development of computer programming concepts and problem solving abilities among ten-year-olds learning LOGO*. Unpublished doctoral dissertation, Syracuse University, Syracuse, New York.
- Statz, J., Folk, M., & Seidman, R. (1973). *Final report: Syracuse LOGO project*.
- Sylva, K., Bruner, J., & Genova, P. (1976). The role of play in the problem-solving of children 3- to 5-years old. In J. Bruner, A. Jolley, & K. Sylva (Eds.), *Play*. New York: Basic.
- Thorndike, E., & Woodworth, R. (1901). I. The influence of improvement in one mental function upon the efficiency of other functions; II. The estimation of magnitudes; III. Functions involving attention, observation and discrimination. *Psychological Review*, 8, 247-261, 384-395, 553-564; cited in Ellis (1965).
- Thorndyke, P., & Stasz, C. (1980). Individual differences in procedures for knowledge acquisition from maps. *Cognitive Psychology*, 12, 137-175.
- Weir, S., & Emanuel, R. (1976, January). Using LOGO to catalyse communication in an autistic child. *D.A.I. Research Report No. 15*. University of Edinburgh.
- Watt, D. (1982). LOGO in the schools. *Byte*, 7, 116-134.
- Woodworth, R., & Schlosberg, H. (1954). *Experimental psychology*. New York: Holt, Rinehart & Winston.

ESSAY REVIEWS

II

SELF-FULFILLING PROPHECIES IN THE CLASSROOM: DEAD END OR PROMISING BEGINNING?

Valerie S. Smead

Western Illinois University

Teacher expectancy research has been grinding along since the initial burst of enthusiasm in the late sixties and early seventies. Laboratory studies artificially manipulating teachers' expectations still appear (Kedar-Voivodas & Tannenbaum, 1979; Taylor, 1979; Ysseldyke & Foster, 1978). Naturalistic studies continue (Good, Cooper, & Blakey, 1980; Luce & Hoge, 1978). Continued scrutiny also reveals further methodological and theoretical problems (Dusek, 1975; Reschly & Lamprecht, 1979; Rogosa, 1980; West & Anderson, 1976).

Originally, such research was in tune with the sociopolitical climate of the time. *Pygmalion in the Classroom* (Rosenthal & Jacobson, 1968) appeared during a period of liberalism and political unrest. Helping efforts at such times tend toward adjusting social institutions to meet individual needs (Levine & Levine, 1970). Broad change is seen as desirable and achievable. The notion that teacher expectations were responsible for students' differential achievement outcomes fit in well. If the teacher got what he or she expected, rather than what the students were capable of producing, social injustice was being perpetuated in part by the educational system. Teachers expecting certain children to do well (perhaps white or upper SES children) or badly (perhaps minority or lower SES children) were producing what they anticipated in the absence of real differences initially. Teachers and school systems, not children, were in error.

Now the temper of the time has changed. According to a recent *Reader's Digest* article, "Why Kids Can't Write," teachers are referring to the last decade as the "Bad Period" in education (Brandt, 1981). Considerably milder interpretations of

Dr. Smead is an Assistant Professor of Psychology at Western Illinois University, where she is also a staff member of the Psychology Clinic. Her teaching and research areas are in community mental health, applied psychology, and human development.

teacher expectations as sustaining rather than causing student achievement differences are also beginning to appear (Cooper, 1979; Good et al., 1980). While not minimizing the significance or need for social concern implied, Cooper (1979), in reviewing the literature, maintains:

For now, then, it seems best to conclude that expectations influence performance, but they likely sustain it at a preexisting level or allow innate differences in student performance to emerge rather than radically alter its course. (p. 392)

It will be argued here that despite methodological difficulties, competing rival hypotheses, and a different political climate, self-fulfilling prophecies in the classroom are still potentially important phenomenon deserving further study. Based on a review of selected research, an interactional framework for conceptualizing the effect will be presented. It does not cast teachers or the educational system as villains, nor students as passive victims. Rather, the interaction over time between teachers, students, and educational system is used as the focus for investigation.

Interactional approaches are not new in conceptualizing the teacher expectancy effect (see Braun, 1976; Cooper, 1979; Finn, 1972; Good, 1980) or in educational research in general (see Hobbs, 1975; Smead, 1977; Tracy, Gibbins, Kladder, & Daggy, 1974; Winne & Marx, 1977). Perhaps this particular effort will prove useful because it attempts to acknowledge and build upon the complex interactions which have been implicated in expectancy research. Ongoing reciprocal relationships between teacher and student are not seen as methodological dilemmas, but as the focus for investigation.

First, the research will be reviewed generally to provide an historical perspective which highlights some of the difficulties that researchers have confronted. Next, findings will be presented within an interactional framework. Both teacher and child effects on the relationship between expectations, behavior, and achievement will be considered. Other expectations in the classroom will also be addressed. This will be followed by speculations on stalking the expectancy effect over the course of a child's school career.

Strengths and Weaknesses of the Current Data Base

Though there have been instances where meager or negative results have been reported (Fleming & Anttonen, 1971; José & Cody, 1971; Humphreys & Stubbs, 1977; Luce & Hoge, 1978), many studies targeted towards various aspects of expectancy transmission have reported significant effects. Investigations utilizing videotapes (Foster, Ysseldyke, & Reese, 1975; Mason & Larimore, 1974; Ysseldyke & Foster, 1978), hypothetical instances (Bar-tal & Saxe, 1979), short term laboratory teaching situations (Beez, 1970), and "phantom" students (Taylor, 1979) appear to have a particularly good batting average. In addition, more naturalistic studies have reported correlational relationships between teacher expectations and student achievement or ability in actual classroom situations (Brophy & Good, 1970; Dusek & O'Connell, 1973; O'Connell, Dusek, & Wheeler, 1974; Palardy, 1969; Seaver, 1973; Willis, 1973). Such studies have also led to documentation of links between teacher expectations and teacher and student behavior (Brophy & Good, 1970; Good et al., 1980). In fact, when all the studies of expectation (teacher, experimenter, laboratory, field) are combined in a meta-analysis of the expectancy effect, "the reality of the phenomenon is beyond doubt and the mean size of the effect is clearly not trivial" (Rosenthal & Rubin, 1978, p. 385).

Unfortunately, there are problems with interpreting these data which could limit the seemingly impressive implications. Definite strengths in terms of the psycho-

logical credibility of the concept (Braun, 1976) and the amount of research it has stimulated are difficult to deny. Documentation that expectancies can be induced in the laboratory and that naturally occurring expectations relate to behavior and achievement in the field also may be beyond dispute. However, compelling demonstrations that teachers actually bias (see Dusek, 1975; West & Anderson, 1976), students' learning success, due to inaccurate expectations, are not as readily available.

The results of laboratory studies utilizing videotapes, hypothetical instances, phantom students, and even brief interactions are difficult to generalize to the ongoing interactions between teachers and students which occur in school situations. As Mason and Larimore (1974) warned more than a decade ago:

This study is a laboratory situation and teachers may not interact with TV sets and psychology reports in the same way that they interact, with children and psychological reports in real situations. (p. 48)

In some cases such studies are questionable even as laboratory demonstrations (Dusek, 1975; Reschly & Lamprecht, 1979). In the absence of information except expectancy manipulations and brief interaction with a videotape or a stranger, it is not particularly surprising that subjects accept and act on the information which they have received. Which piece they chose is of interest and usually intended to generate hypotheses to be tested against actual teachers in real classrooms. It is not a demonstration that: (1) teachers actually form expectations based on labels, psychological reports, cumulative folders, etc. which override student capacities as demonstrated in their classrooms; (2) such information is necessarily an invalid source for forming expectations; perhaps not all psychological reports and diagnostic labels represent "bogus" information; or (3) in those cases where the information is invalid, that teachers will create prophecy effects through their own credulity. One investigation which extended teacher exposure time to videotaped student expectee beyond the usual 10 to 12 minutes reported that, in fact, the effect dissipated after 30 minutes (Reschly & Lamprecht, 1979). Initially, teachers accepted the expectations they had been given. However, after watching the child for just one half-hour, the induced expectations lost their hold.

Studies in the field utilizing naturally occurring teacher expectations bypass some of the difficulties described above. They are based upon ample interaction and actual classroom behavior. However, correlational relationships between teacher expectation and student ability or achievement are subject to another difficulty in interpretation. As Crano & Mellon (1978) suggest, documenting such relationships may "limit the outcome of the research to, at best, a statement of the obvious" (p. 40). That is, competent teachers would be expected to accurately forecast student achievement. A significant relationship between expectation and achievement, therefore, may reflect an expectancy effect, or astute teachers, or some combination of the two (see Crano & Mellon, 1978; Dusek, 1975; West & Anderson, 1976).

Attempts to address this issue have yielded uneven results and further methodological difficulties. One approach involves artificial manipulation of teacher expectation, as did the original Rosenthal and Jacobson study. Such a move beyond correlation to experimental control, though appealing, again raises questions concerning generalizability. The demand characteristics of supposedly correct data, supplied by supposedly knowledgeable and honorable researchers might, on occasion, override teachers' better judgments for a time. Such situations do not typify day-to-day life in the school. In some cases, significant effects have occurred

through manipulation of expectation (Meichenbaum, Bowers, & Ross, 1969; Zanna, Sheras, Cooper, & Shaw, 1975). In others, teachers have indicated that they did not, in fact, find manipulated test results credible (Fleming & Anttonen, 1971; José & Cody, 1971). In one instance where naturally occurring teacher expectations were compared to an attempt at manipulation, only those occurring naturally related to student achievement (Dusek & O'Connell, 1973). This correlational relationship held on three different testing occasions during the school year (Dusek & O'Connell, 1973) and even after the children had gone on to new classrooms the next year (O'Connell, Dusek, & Wheeler, 1974). For further critiques of artificial manipulation see Dusek (1975), Finn (1972), Pelligrini and Hicks (1972).

Another tack involves detecting inappropriate expectations and relating them to subsequent achievement (Palardy, 1969; Sutherland & Goldschmid, 1974). However, with young students without a history of test scores, grades, etc., making an accurate determination of ability, against which expectations may be evaluated, is difficult. It is questionable whether a pre-reading screening test (Palardy, 1969) or even four WISC subtests and a group intelligence test (Sutherland & Goldschmid, 1974) given early in elementary school are sufficient bases for determining the accuracy of teacher expectations. The significant effects which emerged may document prophecy effects for inaccurate teacher expectations, or they may document that teachers were, in actuality, better predictors of future performance than the measuring devices utilized.

A final approach involves utilizing cross-lagged panel designs in an attempt to determine causal predominance. The question addressed is, are expectations more likely to be accurate reflections of previous achievement or may achievement more appropriately be considered a reflection of previous expectations? Proponents of this technique maintain that with two correlated variables such as expectation and achievement, this procedure allows for determination of the direction of causality. Thus far, in one instance teacher expectation was supported as a causal variable in terms of elementary students' achievement (Crano & Mellon, 1978); in another such support was not obtained in terms of high school students' grades (Humphreys & Stubbs, 1977). In addition, some maintain that cross-lagged panel designs are totally inadequate methods for attempting to unravel causality and should be abandoned (Rogosa, 1980). Utilizers of this technique do acknowledge that if a third variable were operating on both variables to unequal extents and went by undetected, predominance would inappropriately be assigned to one of the original variables (Crano & Mellon, 1978; Humphreys & Stubbs, 1977).

In sum, despite an impressive array of significant results, it is difficult to demonstrate conclusively that teacher expectations bias student learning. Expectations which have been artificially manipulated in the laboratory or in actual school situations are subject to questions concerning generalizability. At times, manipulations may not be believed by teachers, who may be operating on the basis of their own naturally occurring expectations. Credibility issues are becoming more important in the laboratory, as well. At the end of a recent experiment more than half of the subjects indicated the suspicion that it related in some way to teacher expectations (Taylor, 1979). Studies which forego experimental control through manipulation are confronted with the problem that naturally occurring expectations may be accurate. If such is the case then teacher expectations are not biasing student learning, but rather correctly forecasting it.

Expectations in an Interactional Framework

Starting with the oft-documented correlation between teacher expectations and student achievement, a different question might be asked. Rather than further attempts to determine how much expectations influence (rather than are influenced by) student ability and achievement, a bi-directional relationship might simply be accepted. In addition to the appeals of common sense and compromise, until totally accurate measures of child academic potential are devised for preschool children, teacher expectation cannot be separated from student achievement and potential. That is, once exposed to teachers who behave in line with their expectations, or alternatively, expect in line with their behaviors (see Brophy & Good, 1970; Good et al., 1980; West & Anderson, 1976), the effects of such treatment on standardized test performance cannot be discounted. High teacher expectancy students may learn more. Unless ability is seen as a fixed attribute, they may even increase their potential for learning in contrast to lower expectancy students. Over the course of a child's school career, in some cases, expectation, behavior, and achievement, therefore, might be expected to follow upward or downward spirals. Once a child has entered kindergarten, teacher expectations and behaviors may be reflected in his or her learning success, grades, and standardized test performance which, in turn, may influence the next teacher's expectations and behaviors (see Finn, 1972, and Good, 1980, for somewhat similar discussions).

From this perspective, instead of further attempts to unravel these variables, attempts to chart, understand, and eventually change the course of their covariation during a child's school career are suggested. The question becomes why, for some children, do expectations and achievements begin a downward spiral and how may such a spiral be prevented.

In addressing this question, several additional areas of expectancy research must be considered. First, academic expectations and behavior in the classroom will be discussed. Student as well as teacher expectations and behaviors have surfaced as meaningful variables here. Secondly, teachers and students may have expectations which go beyond the assumed academic potential of the student. Finally, time and situation may influence the interaction between expectations, behavior, and achievement. Each will be discussed in turn.

Expectations and Behavior

The link between expectations and classroom behavior has received intensive investigation (see Brophy & Good, 1970; Cooper, 1979; Good, 1980; Good et al., 1980). As mentioned, several teaching behaviors which covary with teachers' expectations have been uncovered. For instance, in some cases, teachers wait less time for low expectancy students' responses, demand less from them, criticize more, and seat them further from the teacher's desk. Good et al. (1980) report on these and seven other differential teaching behaviors which have been reported in at least two studies each. Such differences do not, however, occur for all teachers. Based on a large pool of observational studies done in one geographic region, Good (1980) estimates that approximately one-third of the teachers acted in ways which could have exaggerated differences between high and low students unnecessarily.

At this point students must also be considered. Do "good" and "bad" students treat teachers differently? Do children see themselves as good and bad students and act on their own expectations? The answer may be affirmative on both counts. More than a decade ago in a correlational study utilizing first graders, it was re-

ported that "the highs seek out the teacher and initiate interactions with her more frequently than the lows . . . the highs much more frequently show their work . . . or ask questions about it" (Brophy & Good, 1970, p. 368). This possibility of student as active agent in the expectancy drama was initially pursued by only a few (Zanna et al., 1975). Perhaps these few were impervious to the tenor of the times, or perhaps aware earlier than many that the position of passive victim is no more comfortable than that of villain. One early attempt at theory building included student expectation as a central variable (Finn, 1972). Several studies manipulated student as well as teacher expectation (Kern, 1973; Mulligan, 1973; Rappaport & Rappaport, 1975; Zanna et al., 1975). In the latter three studies, student expectation appeared to have either a similar or greater impact than that of the teacher.

More data have since accumulated. Students' naturally occurring expectations for their achievement in eighth grade mathematics class correlated with later achievement across two dependent measures and times of the year (Smead & Chase, 1981). Learning disabled students, it now appears, are operating within what Finn (1972) suggested was a network of expectations. The students, their mothers, and their teachers all hold low expectations for their future achievement; students themselves may react to success and failure in ways which facilitate the occurrence of such expectations (Bryan & Pearl, 1979). Even among regular second and third graders, those whom the teacher rated as being in the bottom third of the class academically rated themselves significantly lower than did those in the top third (Stipek, 1981). These self-ratings also correlated with ratings they received from their peers. Similarly, eighth grade mathematics students, particularly high expectancy students, reported substantial agreement between their own expectations and those they felt that their parents and peers held for them (Smead & Chase, 1981).

Finally, additional research confirms that students for whom teachers hold high and low expectations behave differently in the classroom. It is introduced with the caution:

It must be kept in mind that the data to be reported are correlational. It is impossible to determine, therefore, if differences in the classroom interaction are caused by teacher or student behavior (or if influence is bidirectional). (Good et al., 1980, p. 380)

Thus, child as well as teacher effects may be found in the ongoing interaction between expectations, behavior, and achievement.

Beyond Academic Potential

Cooper (1979) has argued that teachers' differential behaviors toward high and low expectancy students might reflect beliefs about how controllable such children are. For instance, criticizing low expectancy students more and praising them less might serve to discourage low control exchanges between the teacher and these students (see Cooper, 1979). Other work suggests that not only may achievement expectations contain associated beliefs about controllability, but that teachers and perhaps students may have expectations which are not based only upon assumed academic potential of the student. These expectations may influence such potential, however.

For instance, two types of expectations were measured in the Crano and Mellon (1978) study. In addition to teachers' expectations for students' academic performance, expectations for social performance (in terms of disobedience, pleasantness to have in class, and attitude toward school work, including work habits) were obtained. They reported that the "most positive and consistent results" (p. 47) in

terms of expectations impact upon later academic achievement, were obtained from these social expectations and concluded:

It is obvious . . . that the effects of social evaluations and expectations, based as they are on aspects of a child totally incidental to intellectual promise, represent a potential educational inequity of even greater import than that emphasized by Rosenthal and Jacobson. (1968, p. 48)

Elsewhere, factor analytic study of teacher expectations for a variety of child characteristics (conduct, school performance, sociability) suggested a first factor that was fairly similar for all teachers (Garner & Bing, 1973). This factor included attitudes toward work, school performance and intellectual potential. Interestingly, the positive aspects of many other dimensions loaded on this factor, "suggesting a fairly widespread stereotype of the 'good' pupil, the 'average' pupil, and the 'poor' pupil" (p. 36). In this instance a consistent relationship between expectations and verbal interchange, the dependent measure utilized, was not reported. Elsewhere, teacher rating of student motivation correlated significantly with a variety of teacher behaviors while teacher ratings of ability and achievement generally did not (Luce & Hoge, 1978). In response to hypothetical instances (Kedar-Voivodas & Tannenbaum, 1979), teachers were more influenced by descriptions of classroom behaviors than by sex, label, and therapy condition, each of which had no significant effect. In addition, acting out behaviors had more of a detrimental effect on teacher expectation than did withdrawn behaviors of equal severity.

Students, themselves, have expressed the belief, again in response to hypothetical instances, that student effort is more critical for teacher judgment than student ability (Harari & Covington, 1981). This view held for primary and secondary students. In the primary grades, the reputation of being a hard worker was also a student value. Based on further exploration with semistructured interviews, students in grades 1, 2, and 4 appeared to see effort and outcome as synonymous. Those who tried hard were expected to do well; in fact, they could even improve their ability. In actual classroom situations, early elementary students justified their judgments of peers' ability most frequently by referring to their work habits (Stipek, 1981).

In sum, it appears that child motivation, manageability, and general willingness to play the good student role may also influence achievement expectations and/or behaviors in the classroom.

Expectation, Achievement, and Behavior Across Time and Situation

In attempting to chart the covariation of expectations, behavior, and achievement over the course of a school career, not only must differential teacher expectation and behavior be taken into consideration, but children, themselves, cannot be ignored. Their behavior, expectations and potential for playing the good student role may also contribute.

As discussed, once children enter kindergarten, the effects of teacher expectations and behaviors may be reflected in childrens' learning, grades, and standardized test performance. At this point, it appears that the teacher might also influence student behaviors and expectations which, again, might be carried into subsequent classrooms. However, the teacher is no longer clearly the starting point. Child differences in behavior may also cause teacher differential behaviors and perceptions. Child expectations, motivation, and manageability also may not be entirely contingent on teacher treatment. Therefore, some mix of teacher and child effects

on achievement and potential may be carried into new classrooms and aid in perpetuating upward and downward spirals. The issue which must be addressed is who and what is more important at different points in the child's school career. Unfortunately, with interactional data such as these, considerable speculation is involved in attempting to chart expectations across the course of a school career.

Teacher expectations and behavior may be more important starting with the early grades of elementary school than they will be later. Several investigators have mentioned this as a potentially critical time (Cooper, 1979; Good, 1980; Sutherland & Goldschmid, 1974). For many elementary school children the teacher's credibility rating is high; she or he is the one who "knows" (Braun, 1976). Also, at this age children are more impressionable and anxious to please adults (Good, 1980). Several studies documenting relationships between teacher expectation and student achievement, in fact, have involved these years (Brophy & Good, 1970; Crano & Mellon, 1978; Dusek & O'Connell, 1973; O'Connell, Dusek, & Wheeler, 1974; Palardy, 1969; Seaver, 1973). Longitudinal data across the first four grades, an interesting byproduct of one of the cross-lagged panel designs (Crano & Mellon, 1978) also suggest that during this period expectations may tend toward increasing consistency between teachers. In particular, the following correlations were obtained for teacher expectations (academic and social) .289 between teacher, grade one, and teacher, grade two; .550 between two and three; and .562 between three and four. Finally, West and Anderson (1976) have suggested that teacher expectation would be more effective when previous interaction is minimal, the content area is new and not too similar to previously learned material, and teacher behavior is essential for gaining knowledge. For many children, early elementary school may meet these criteria.

Downward or upward spirals need not start early in school, however, and if they do the teacher need not be the only active participant. It has been repeatedly documented that in some classrooms, teachers do not appear as expectancy prone as in others (Brophy & Good, 1970; Good et al., 1980; Good, 1980). Therefore, children might escape teachers who manifest differential behaviors toward high and low students. In addition, they might escape relegation to high and low categories. In one instance involving older students (6th to 8th graders), teachers indicated more confidence in their estimates of students who scored toward the extreme ends of the continuum (Felson & Bohrnstedt, 1980). This relationship was noted particularly with positive extremes. Thus, even in the case of expectancy prone teachers, if student performance is not extreme, but more towards the middle of the continuum, then differential expectations might be precluded.

Even at this age, child, parent, and peer contributions cannot be discounted. In extreme cases, second and third grade students may already manifest differential expectations for themselves (Stipek, 1981). Perhaps students who adopt behaviors associated with motivation, manageability, and the good student role may prove less susceptible to negative spirals than peers of similar ability who do not. Expectancy resistant parents might be another lever—they might rebut low grades, transmit essential knowledge, or mold good student behaviors themselves when child or teacher is in danger of initiating such a spiral. Peers who model appropriate behavior and attitudes and provide feedback based on work habits may also contribute.

As time goes on, teacher effect may be further diluted as children's own expectations take on more importance at the later elementary or secondary level. First, classroom behavior might be expected to be "progressively determined by student's

past experiences and performances in similar situations" (Good, 1980, p. 117). Secondly, as discussed below, views of their own abilities may be differentiating and solidifying.

Kifer (Note 1 reported by Newman & Newman, 1981) asked youngsters to compare their school ability to that of other children in their grades. Responses of A and B students in grades two through eight were compared to those of D and E students. In the early grades estimates of ability were not related to grades; the trend was for all students to express high self-concepts of ability. But as grade level increased there was a steady downward trend in expressed self-concept of ability for D and E students. By grade eight they were expressing quite low appraisals of their ability while A and B students continued to express high appraisals. Elsewhere, the validity of students' expectations (based on specific performance tasks) was found to increase across first, third, and fifth graders; even at the fifth grade level validity did not exceed 1.8 on a 3-point scale (Clifford, 1975). Overestimation tendencies were reported as was failure to process task feedback.

Thus, not only may behaviors and attitudes towards school crystallize with time, but children may enter new classrooms with progressively more clear-cut expectations for themselves. They may also be changing their earlier view that "ability as a concept is fluid, malleable, and highly dependent on effort" (Harari & Covington, 1981, p. 25). In that investigation, by the sixth grade students no longer expressed the idea that effort causes ability change. By eighth grade ability was seen as a more stable factor and as a "strong moderator of the effort-outcome covariation" (p. 25). The view that anyone, irrespective of ability, who studies hard will do well was no longer endorsed. By eleventh grade outcome was seen primarily as a function of ability.

Finally, on entering junior high school or the later elementary grades in some schools, students are exposed to not one but several teachers. Many times these teachers are exposed to not 30 but over one hundred students. Often teacher behavior is no longer essential to accumulating knowledge—other sources are homework, reading the text, or copying a friend's notes. The content area may not be novel, and previously formed expectations for reading, mathematics, and even physical education may have direct bearing on current student performance in the classroom.

However, just as early elementary teachers' expectations and behaviors are not seen as omnipotent, neither are those of secondary students. Not only are new peer groups who express different expectations and attitudes possible, but so are teachers or subjects which suddenly inspire them. Students who might be most flexible at this stage are those who have not formed stable images of themselves as learners due to inconsistent performance or attitudes toward school. Also of interest is that group in the broad middle range of achievement who can acquire sufficient skills to "take off" if appropriately motivated from within or without.

Conclusion

The question is not whether expectations bias achievement. This review suggests that child and teacher expectations, behavior, and child achievement work together over the course of a child's school career. The question is which variables are most accessible to modification at which point in time. At this point, much speculation is involved in charting what might be called the expectancy-behavior-achievement effect. Longitudinal studies monitoring these variables from

kindergarten through secondary school are needed. At present, the possibility of a gradually tightening web of variables is suggested for some children. Those who do not play the good student role and score towards the lower end of the continuum may be among those particularly vulnerable to negative spirals.

Also suggested are multiple intervention points from child expectation, to teacher preference for particular behaviors, to preparing expectancy resistant parents. In the years since *Pygmalion in the Classroom*, the magical connotation of teacher expectations has disappeared. Now, at times, all involved appear to be potential "victims" (see Ryan, 1971, 1976)—teachers, parents, peers, and students who come to believe they are dumb. All appear to have some potential power in halting downward spirals, as well. Further experimental manipulations are necessary. However, rather than continuing ill-fated attempts to untie variables which naturally work together, these must address the impact of particular teacher, parent, or student-centered interventions at particular times in the child's school career.

Notes

1. Kifer, E. The impact of schooling on perceptions of self. Paper presented at the self concept symposium, Boston, September, 1978.

References

- Bar-tal, D., & Saxe, L. (1979). Teachers' information processing: Effect of information about pupils on teachers' expectations and affect. *Psychological Reports*, 44, 599-602.
- Beez, W. V. (1970). *Influence of biased psychological reports on "teacher" behavior and pupil performance*. Unpublished doctoral dissertation, Indiana University.
- Brandt, A. (1981). Why kids can't write. *Reader's Digest*, 118, 99-102.
- Braun, C. (1976). Teacher expectation: Sociopsychological dynamics. *Review of Educational Research*, 46, 185-213.
- Brophy, J. E., & Good, T. C. (1970). Teachers' communication of differential expectations for children's classroom performance: Some behavioral data. *Journal of Educational Psychology*, 61, 365-374.
- Bryan, T., & Pearl, R. (1979). Self concepts and locus of control of learning disabled children. *Journal of Clinical Child Psychology*, 223-226.
- Clifford, M. M. (1975). Validity of expectation—A developmental function. *Alberta Journal of Educational Research*, 21, 11-17.
- Cooper, H. M. (1979). Pygmalion grows up: A model for teacher expectation communication and performance influence. *Review of Educational Research*, 49, 389-410.
- Crano, W. D., & Mellon, P. M. (1978). Causal influences of teachers' expectations on children's academic performance: A cross-lagged panel analysis. *Journal of Educational Psychology*, 70, 39-49.
- Dusek, J. B. (1975). Do teachers bias children's learning? *Review of Educational Research*, 45, 661-684.
- Dusek, J. B., & O'Connell, E. J. (1973). Teacher expectancy effects on the achievement test performance of elementary school children. *Journal of Educational Psychology*, 65, 371-377.
- Finn, J. D. (1972). Expectations and the educational environment. *Review of Educational Research*, 42, 387-410.
- Felson, R. B., & Bohrnstedt, G. W. (1980). Attributions of ability and motivation in a natural setting. *Journal of Personality and Social Psychology*, 39, 799-805.

- Fleming, E. S., & Anttonen, R. G. (1971). Teacher expectancy as related to the academic and personal growth of primary-age children. *Monographs of the Society of Research in Child Development*, 36, Serial No. 145.
- Foster, G. G., Ysseldyke, J. E., & Reese, J. H. (1975). I wouldn't have seen it if I hadn't believed it. *Exceptional Children*, 41, 469-473.
- Garner, J., & Bing, M. (1973). The elusiveness of Pygmalion and differences in teacher-pupil contacts. *Interchange*, 4, 34-42.
- Good, T. L. (1980). Classroom expectations: Teacher and pupil interactions. In J. H. McMillan (Ed.), *The social psychology of school learning*. New York: Academic Press.
- Good, T. L., Cooper, H. M., & Blakey, S. L. (1980). Classroom interaction as a function of teacher expectations, student sex, and time of year. *Journal of Educational Psychology*, 72, 378-385.
- Harari, O., & Covington, M. V. (1981). Reactions to achievement behavior from a teacher and student perspective: A developmental analysis. *American Educational Research Journal*, 18, 15-28.
- Hobbs, N. (1975). *The futures of children*. San Francisco: Jossey-Bass.
- Humphreys, L. G., & Stubbs, J. (1977). A longitudinal analysis of teacher expectation, student expectation, and student achievement. *Journal of Educational Measurement*, 14, 261-270.
- José, J., & Cody, J. (1971). Teacher-pupil interaction as it relates to attempted changes in teacher expectancy of academic ability and achievement. *American Educational Research Journal*, 8, 39-49.
- Kedar-Voivodas, G., & Tannenbaum, A. J. (1979). Teachers' attitudes toward young deviant children. *Journal of Educational Psychology*, 71, 800-808.
- Kern, J. C. (1973). Experimenter effects in pupil testing (Doctoral dissertation, The City University of New York, 1972). *Dissertation Abstracts International*, 33, 3919B.
- Levine, M., & Levine, A. (1970). *A social history of helping services*. New York: Appleton-Century-Crofts.
- Luce, S. R., & Hoge, R. D. (1978). Relations among teacher rankings, pupil-teacher interactions, and academic achievement: A test of the teacher expectancy hypothesis. *American Educational Research Journal*, 15, 489-500.
- Mason, E. J., & Larimore, D. L. (1974). Effects of biased psychological reports on two types of teachers' ratings. *Journal of School Psychology*, 12, 46-50.
- Meichenbaum, D. H., Bowers, K. W., & Ross, R. R. (1969). *Journal of Personality and Social Psychology*, 13, 306-316.
- Mulligan, J. P. (1973). Teacher and student expectations and student achievement in college reading improvement courses (Doctoral dissertation, Rutgers University, 1972). *Dissertation Abstracts International*, 33, 3776A.
- Newman, P. R., & Newman, B. M. (1981). *Living: The process of adjustment*. Homewood: Dorsey.
- O'Connell, E. J., Dusek, J. B., & Wheeler, R. J. (1974). A follow-up study of teacher expectancy effects. *Journal of Educational Psychology*, 66, 325-328.
- Palardy, J. M. (1969). What teachers believe—What children achieve. *Elementary School Journal*, 69, 370-374.
- Pellegrini, R. J., & Hicks, R. A. (1972). Prophecy effects and tutorial instruction for the disadvantaged child. *American Educational Research Journal*, 9, 413-418.
- Rappaport, M. M., & Rappaport, H. (1975). The other half of the expectancy equation: Pygmalion. *Journal of Educational Psychology*, 67, 531-536.
- Reschly, D. J., & Lamprecht, M. J. (1979). Expectancy effect of labels: Fact or artifact? *Exceptional Children*, 45, 55-58.
- Rogosa, D. A. (1980). A critique of cross-lagged correlation. *Psychology Bulletin*, 88, 245-258.

- Rosenthal, R., & Jacobson, L. (1968). *Pygmalion in the classroom*. New York: Holt, Rinehart & Winston.
- Rosenthal, R., & Rubin, D. B. (1978). Interpersonal expectancy effects: The first 345 studies. *The Behavioral and Brain Sciences*, 3, 377-415.
- Ryan, W. (1971, 1976). *Blaming the victim*. New York: Vintage Books.
- Seaver, W. B. (1973). Effects of naturally induced teacher expectancies. *Journal of Personality and Social Psychology*, 28, 333-342.
- Smead, V. S. (1977). Ability training and task analysis in diagnostic/prescriptive teaching. *The Journal of Special Education*, 11, 113-125.
- Smead, V. S., & Chase, C. I. (1981). Student expectations as they relate to achievement in eighth grade mathematics. *Journal of Educational Research*, 2, 115-120.
- Stipek, D. J. (1981). Children's perceptions of their own and their classmates' ability. *Journal of Educational Psychology*, 73, 404-410.
- Sutherland, A., & Goldschmid, M. L. (1974). Negative teacher expectation and IQ change in children with superior intellectual potential. *Child Development*, 45, 852-856.
- Taylor, M. C. (1979). Race, sex, and the expression of self-fulfilling prophecies in a laboratory teaching situation. *Journal of Personality and Social Psychology*, 37, 897-912.
- Tracy, M. L., Gibbins, S., Kladder, F. W., & Daggy, N. (1975). *Case conference: A simulation and source book*. Ann Arbor: Distributed by the Publication and Distribution Service of the University of Michigan Press.
- West, C. K., & Anderson, T. H. (1976). The question of preponderant causation in teacher expectancy research. *Review of Educational Research*, 46, 613-630.
- Willis, S. L. (1973). Formation of teachers' expectations of student academic performance (Doctoral dissertation, University of Texas, Austin, Texas, 1972). *Dissertation Abstracts International*, 33, 4960A.
- Winne, P. H., & Marx, R. W. (1977). Reconceptualizing research on teaching. *Journal of Educational Psychology*, 69, 668-678.
- Ysseldyke, J. E., & Foster, G. G. (1978). Bias in teachers' observations of emotionally disturbed and learning disabled children. *Exceptional Children*, 44, 613-615.
- Zanna, M. P., Sheras, P. L., Cooper, J., & Shaw, C. (1975). Pygmalion and Galatea: The interactive effect of teacher and student expectancies. *Journal of Experimental Social Psychology*, 2, 279-287.

PREPARATION OF MANUSCRIPTS

1. All manuscripts must be typewritten, double spaced, and submitted in duplicate. An abstract of approximately 100 words in length, typed on a separate page, should be provided.
2. Tables must be numbered in Arabic numerals with the word 'Table' centered and in capital letters, e.g., TABLE 4. The heading of the table is to be centered below and typed in capitals. The format of tables should conform to the specifications in the APA Publications Manual.
3. Graphs and charts should be used only if essential. They must be carefully prepared on separate sheets in India ink, ready for reproduction. Graphs must be properly labelled using Arabic numerals, e.g., Figure 3.
4. Each table or figure should be presented on a separate page. The position of tables and graphs should be clearly indicated within the text by inserting at the relevant point the phrase (Insert Table 2 here).
5. References should appear in parentheses following the reference citing the author's name (unless the name appears in the text), the year of publication, and page number if appropriate. For direct quotations, the reference should be cited and the page number given in brackets before the final punctuation of the quotation. The references should be listed alphabetically by authors' last names at the end of the manuscript under the heading, *References*.
6. Explanatory notes, numbered consecutively and identified in the text with a superscript, may be included under the heading of *Notes*. They should be double spaced and placed at the end of the manuscript immediately preceding the *References*. The citing of references and quotations in the *Notes* should conform to the procedures outlined in No. 5 above.
7. Spelling shall conform to the *Oxford English Dictionary* or to the *Dictionary of Canadian English*, except in the case of direct quotations which must conform to the original. Editorial alterations will be made if necessary.
8. In matters of style, the APA Publications Manual is considered definitive.

